UNIVERSIDADE DE LISBOA

INSTITUTO DE EDUCAÇÃO

O CONHECIMENTO DIDÁTICO EM ESTATÍSTICA DE DUAS PROFESSORAS DO ENSINO SECUNDÁRIO A PARTIR DAS SUAS PRÁTICAS

Sandra Maria Oliveira Quintas

Orientadores: Prof.ª Doutora Hélia Margarida Aparício Pintão de Oliveira
Prof.ª Doutora Rosa Antónia de Oliveira Figueiredo Tomás Ferreira

Tese especialmente elaborada para a obtenção do grau de Doutor em Educação, na área de especialidade de Didática da Matemática

2017

UNIVERSIDADE DE LISBOA

INSTITUTO DE EDUCAÇÃO

O CONHECIMENTO DIDÁTICO EM ESTATÍSTICA DE DUAS PROFESSORAS DO ENSINO SECUNDÁRIO A PARTIR DAS SUAS PRÁTICAS

Sandra Maria Oliveira Quintas

Orientadores: Prof.ª Doutora Hélia Margarida Aparício Pintão de Oliveira
Prof.ª Doutora Rosa Antónia de Oliveira Figueiredo Tomás Ferreira

Tese especialmente elaborada para a obtenção do grau de Doutor em Educação, na área de especialidade de Didática da Matemática

Júri:

Presidente: Doutor João Pedro Mendes da Ponte, Professor Catedrático e Presidente do Conselho Científico do Instituto de Educação da Universidade de Lisboa;

Vogais:

- Doutora Rosa Antónia de Oliveira Figueiredo Tomás Ferreira, Professora Auxiliar da Faculdade de Ciências da Universidade do Porto, orientadora;
- Doutor José António da Silva Fernandes, Professor Associado do Instituto de Educação da Universidade do Minho;
- Doutora Ana Paula Canavarro Teixeira, Professora Auxiliar da Escola de Ciências Sociais da Universidade de Évora;
- Doutora Leonor de Almeida Domingues dos Santos, Professora Associada com Agregação do Instituto de Educação da Universidade de Lisboa;
- Doutora Ana Cláudia Correia Batalha Henriques, Professora Auxiliar do Instituto de Educação da Universidade de Lisboa.

2017

RESUMO

Esta investigação tem por objetivo compreender o conhecimento didático em Estatística de duas professoras de Matemática do ensino secundário, a partir da análise das suas práticas, enquanto lecionam a unidade de Estatística na disciplina de Matemática A, no 10.º ano de escolaridade.

O enquadramento teórico desenvolve-se em torno do conhecimento didático do professor, e respetivos domínios, o qual diz respeito a um conhecimento ancorado na prática letiva e que integra a ação e o pensamento do professor. Apresenta uma revisão de literatura acerca do ensino e aprendizagem da Estatística e destaca um conjunto de recomendações com origem na investigação em Educação Estatística relativas ao ensino-aprendizagem dos conceitos estatísticos centrais no contexto do ensino secundário, tais como variabilidade, população e amostra, medidas de localização e de dispersão, e dados bivariados e regressão linear.

Esta investigação é de natureza qualitativa e segue o paradigma interpretativo, sendo a modalidade escolhida, o estudo de caso. Foram constituídos dois casos de duas professoras de Matemática do ensino secundário, Estela e Lia, com recurso a diversos instrumentos de recolha de dados, entre os quais observações das aulas da unidade de Estatística, entrevistas semiestruturadas de curta duração (antes e após cada aula) e de longa duração, notas de campo, e recolha documental. A análise dos dados é descritiva e interpretativa, tendo sido orientada pelo quadro teórico relativo ao conhecimento didático do professor em Estatística.

Os resultados da investigação evidenciam um conjunto de conhecimentos das professoras relativos a cada um dos quatro domínios do conhecimento didático em Estatística, nos diferentes tópicos do programa, bem como alguns conhecimentos relativos à articulação entre tópicos estatísticos abordados. Surge como implicação para a formação inicial e contínua que esses resultados podem constituir um bom ponto de partida para a construção e desenvolvimento de situações a serem contempladas em contextos formativos na área da didática da Estatística.

Palavras-chave: Conhecimento didático em Estatística, Estatística, Ensino Secundário, Prática do professor.

ABSTRACT

This investigation aims to understand the didactical knowledge in Statistics of two secondary mathematics teachers, through the analysis of their teaching practices, when they teach the Statistics Unit in a 10^{th} grade Mathematics A course.

The theoretical framework is grounded on the model of teachers' didactical knowledge, comprised of four domains, which perceives the teacher's knowledge as being anchored in the school practice, integrating the teacher's actions and thinking. This framework encompasses a literature review about the teaching and learning of Statistics, while highlighting a set of Statistics Education research recommendations related to the teaching and learning of core statistical concepts at secondary level (such as variability, population and sample, center and variation measures, and bivariate data and linear regression).

This qualitative research follows the interpretative paradigm. Two case studies of two secondary school mathematics teachers, Estela and Lia, were built using several data collection instruments, including: (i) lesson observations of the Statistics Unit, (ii) semi-structured interviews of short (before and after each lesson) and long duration, (iii) field notes, and (iv) documental collection. Data analysis is descriptive and interpretative, guided by the theoretical framework regarding the teacher's didactical knowledge in Statistics.

The results evidence several elements regarding both teachers' knowledge in each of the domains, in all curricular topics, as well as concerning the articulation amongst them. This study has implications for pre- and in-service teacher education, given that its results may constitute a good starting point for the creation and development of situations to be explored in formative contexts in the area of Didactics of Statistics.

Keywords: Didactical knowledge in Statistics, Statistics, Secondary school, Teacher's practice

AGRADECIMENTOS

Às minhas orientadoras, Professora Doutora Hélia Oliveira e Professora Doutora Rosa Antónia Tomás Ferreira, um agradecimento profundo pela atenção e disponibilidade que me deram no decorrer deste trabalho. Por tudo o que aprendi nas nossas reuniões e pelos desafios que me lançaram, todos esses momentos influenciaram o meu desenvolvimento como investigadora. Agradeço, de igual modo, as suas amizades.

Ao Centro de formação, às professoras do estudo e aos alunos das suas respetivas turmas, pela colaboração e disponibilidade.

Aos professores do IEUL, pelas excelentes disciplinas que frequentei na parte curricular do doutoramento. E também pela participação no Projeto *Desenvolver a literacia estatística: Aprendizagem do aluno e formação do professor* (contrato PTDC/CPE-CED/117933/2010) financiado pela FCT – Fundação para a Ciência e Tecnologia.

A todos os meus amigos portugueses e estrangeiros que foram fazendo chegar palavras de encorajamento e o seu carinho. Um agradecimento especial à Elisabete, à Rosa Maria, à Alda, à Lurdinhas, à Maria Elisa, à Josefina e à Ana Sílvia.

A todos os que me enviaram artigos para apoiar este trabalho. Um agradecimento especial ao Leo Goldmakher, à Liz Vivas, ao GianCarlo Urzua, ao Fidel Jimenez e ao Dani Ben-Zvi.

À Madalena Consciência, colega de doutoramento, com quem tive valiosas discussões durante o desenvolvimento desta tese e a quem agradeço a sua amizade.

Aos restantes colegas de doutoramento, pela partilha e encorajamento.

À minha família, em especial, ao meu filho (meu companheirozinho de todos os dias, que está sempre pronto para conversar), à minha mãe (que sempre acreditou no fim desta tese) e ao meu pai (que esperou em silêncio), à minha irmã gémea (que me faz rir), à minha avó materna e à Tita (de quem ainda hoje acato conselhos), pelo apoio, incentivo e amor e a quem dedico esta tese.

O Infante

Aos homens ordenou que navegassem
Sempre mais longe para ver o que havia
E sempre para o sul e para que indagassem
O mar a terra o vento a calmaria
Os povos e os astros
E no desconhecido cada dia entrassem

Sophia de Mello Breyner Andresen

ÍNDICE GERAL

ÍNDICE DE FIGURAS

ÍNDICE DOS QUADROS

ÍNDICE DE ANEXOS

CAPÍTULO 1

INTRODUÇÃO

Neste capítulo, apresenta-se os aspetos tidos como mais relevantes para caracterizar a investigação desenvolvida. Assim, começa-se por apresentar a motivação do estudo, seguida do seu contexto e pertinência. Em seguida, enuncia-se o problema e as questões específicas do estudo. Por último, é descrito o modo como o estudo está organizado.

1.1 Motivação para o estudo

O meu interesse pela Estatística começou sobretudo depois de ter concluído a licenciatura em Matemática – ramo educacional. Na minha opinião, o curso, de um modo geral, foi muito teórico e com poucas aplicações. Tendo isso em conta, quando tive a oportunidade, inscrevi-me num mestrado, numa área relacionada com a Matemática, mais concretamente, em Estatística Aplicada, que iria proporcionar-me uma aprendizagem da Estatística por meio de aplicações a situações diversificadas e concretas da vida real. Esta experiência, que decorreu na *Eastern Michigan University-Ypsilanti*, nos E.U.A., julgo ter sido muito interessante e permitiu-me adquirir uma nova perspetiva desta área, principalmente por ter tomado contacto com a Estatística como ciência dos dados.

Lembro-me, em particular, da disciplina de Análise de Dados que frequentei nesse mestrado, em que a maioria das tarefas propostas consistia em responder a uma questão (ou a questões) de um problema, que abordava uma situação real, através de análises estatísticas de modo a obter conclusões com significado. Um exemplo de uma destas tarefas prendeu-se com um dos maiores parques naturais do mundo, o *Yellowstone*

National Park, situado no Estado de Wyoming, nos E.U.A.. Os visitantes deste parque desejam ver o géiser designado por *Old faithful*[1] a entrar em erupção sem ter que esperar demasiado tempo. Como tal, a tarefa consistia em compreender e prever o intervalo de tempo até à erupção seguinte. Tal como esta, todas as tarefas propostas nesta disciplina, e em outras que envolviam métodos de análise de dados, incluíam dados reais e as suas resoluções assumiam fundamentalmente uma natureza exploratória, exigindo uma constante tomada de decisões.

A resolução das tarefas começava, quase sempre, pela escolha da melhor representação gráfica para apresentar os dados, por vezes, através de várias tentativas, quando se suspeitava da presença de *outliers*, ou até recorrer-se a mudanças na escala original dos dados de modo a visualizar-se outras propriedades dos mesmos. Depois passava-se à escolha do modelo estatístico que melhor se ajustava à situação em estudo, partindo-se sempre dos mais simples, sendo que todas as condições assumidas pelo modelo teriam de ser verificadas e, posteriormente, procedia-se à escrita das conclusões. Nesta escrita, também se incluía o motivo dos métodos escolhidos, das representações utilizadas na análise de dados e ligavam-se os resultados obtidos à situação descrita no problema. Na realização destas tarefas, raramente se podia prescindir da discussão em grupo com os colegas que frequentavam a disciplina, uma vez que a participação na discussão ajudava a fortalecer os argumentos, a perceber se os colegas faziam os mesmos raciocínios e, se não fosse o caso, a tentar perceber a origem dessas diferenças. Esta disciplina tinha como um dos propósitos principais desenvolver nos alunos predisposições para o tipo de abordagem que pode ser mais proveitosa na análise de dados. Após a conclusão do mestrado, passei a refletir muito mais sobre como poderia conjugar estes conhecimentos e experiências que tive na área da Estatística com o ensino da Matemática e da Estatística, em particular.

Apesar de não ter uma carreira profissional muito extensa, no ensino do tema da Estatística, no 3.º ciclo, tenho notado que, de um modo geral, os alunos apresentam dificuldades na compreensão dos conceitos estatísticos. Embora sendo capazes de reproduzir algumas definições destes conceitos, bem como de realizar cálculos e

[1] Pelo facto das erupções deste géiser seguirem um padrão relativamente estável, ele foi designado por *Old faithful* (uma possível tradução: *Velho fiel amigo*).

procedimentos, continuam a mostrar que fazem uma apropriação dos conhecimentos com pouco significado. Por exemplo, revelam dificuldades de interpretação de conceitos, em ir além de uma leitura simples de gráficos, etc. Apesar de os alunos chegarem ao 10.º ano de escolaridade relativamente familiarizados com a maioria dos conceitos abordados no tema da Estatística, por estes já terem sido abordados ao longo do ensino básico, as dificuldades que referi acima estão presentes em muitos alunos. De acordo com a minha experiência, parece-me que na realização de tarefas estatísticas diversificadas nas aulas que incluam, em particular, as de cariz investigativo (Garfield & Ben-Zvi, 2008; Shaughnessy, 2006, 2007), acaba por ser natural que essas e outras dificuldades dos alunos emerjam. Contudo, este contexto de trabalho, também mais desafiante para o professor, mesmo para aqueles que já têm alguma experiência em lidar com análises de dados, pode ser particularmente oportuno para que os alunos superem as suas dificuldades e desenvolvam o seu raciocínio estatístico[2]. Para que ocorra um ensino efetivo da Estatística é fundamental que os professores tenham em conta a especificidade da disciplina (e.g., Cobb & Moore, 1997; Ponte & Fonseca, 2001; Rossman, Chance & Medina, 2006; Shaughnessy, 2006; Snee, 1993; Wild & Pfannkuch, 1999), procurem desenvolver a compreensão dos conceitos e representações de uma forma profunda (e.g., Burril & Biehler, 2011; Garfield & Ben-Zvi, 2008; Shaughnessy, 2007), estejam preparados para desenvolver tarefas estatísticas variadas com diferentes níveis de exigência na aula (e.g., Garfield & Ben-Zvi, 2008; Scheaffer, 2006) e tenham cada vez mais conhecimento das dificuldades dos alunos e das suas origens (e.g., Batanero & Diaz, 2010; Sharma, 2005; Shaughnessy, 2007). O objetivo é, assim, promover o desenvolvimento do raciocínio dos alunos (e.g., Burrill & Biehler, 2011; delMas, 2004; Garfield & Ben-Zvi, 2007), num contexto em que haja espaço para discussão das questões, de modo que a negociação de significados conduza à apropriação dos conhecimentos (Ponte, Boavida, Graça & Abrantes, 1997).

Com este trabalho investigativo pretendi aprofundar e ampliar os meus conhecimentos de como ensinar Estatística, de modo a melhorar a minha prática e dar um contributo positivo para o desenvolvimento da investigação na área do ensino da Estatística. Espero que este trabalho forneça elementos relevantes acerca da natureza do conhecimento que os professores do ensino secundário possuem e também de

[2] A noção de raciocínio estatístico é desenvolvida no capítulo 2 (O ensino e aprendizagem da Estatística).

abordagens que possam sustentar o desenvolvimento do seu conhecimento, no âmbito da Estatística. E que, assim, permita apoiar a formação inicial e contínua de professores de Matemática do ensino secundário, atendendo-se à especificidade da Estatística, em cada tópico, de forma mais eficiente. De uma forma geral, este estudo pretende contribuir para o campo da Educação Estatística, em especial no que diz respeito a uma reflexão aprofundada sobre o conhecimento didático do professor em Estatística e seu desenvolvimento.

1.2 Contexto e pertinência do estudo

A Estatística foi introduzida nos programas escolares do ensino secundário, no nosso país, nos anos sessenta, em simultâneo com a reforma do ensino da Matemática marcada pelo movimento da *Matemática Moderna*. Apesar de explicitamente incluídos nos programas, os conteúdos estatísticos, habitualmente relegados para o fim dos programas, acabavam por não ser lecionados aos alunos por falta de tempo ou de interesse (Branco, 2000). Posteriormente, no início dos anos noventa, numa reforma do sistema educativo, este tema foi introduzido nos 2.º e 3.º ciclos do ensino básico no âmbito da disciplina de Matemática (Ponte & Fonseca, 2001; Ponte, 2002). Em 2007, o tema da Estatística foi incluído no programa do 1.º ciclo do ensino básico. Mais recentemente, com a entrada em vigor de novos programas no ensino secundário, nomeadamente, de Matemática A no 10.º ano, o tema da Estatística deixou de incluir o tópico Dados bivariados e regressão linear, que passou a fazer parte do programa de 11.º ano de Matemática A (MEC, 2014), sendo o único tópico estatístico nesse nível de escolaridade. Este novo programa foi implementado pela primeira no 10.º ano do ano letivo 2015/2016.

O presente trabalho, porém, decorre de um estudo realizado anteriormente, com base no programa de Matemática A de 10.º ano homologado em 2001 (Quintas, Oliveira & Ferreira, 2009, 2010), que era o programa seguido pelas professoras que participaram na investigação, no momento da recolha de dados junto das suas turmas. Nesta tese, assume-se que sempre que se mencionar o programa de Matemática A do 10.º ano está-se a considerar o homologado em 2001, caso contrário, será feita uma referência explícita que se está a referir ao que foi aprovado em 2014 (MEC, 2014).

A Estatística tem sido encarada como uma área favorável ao desenvolvimento de certas capacidades expressas nos currículos de Matemática, tais como: interpretar e intervir no real; formular e resolver problemas; comunicar; manifestar rigor e espírito crítico; contribuir para uma atitude positiva face à Ciência e promover uma cidadania ativa e participativa (ME, 2001). De facto, o programa de Matemática A do ensino secundário destaca, mais especificamente, a importância do tema da Estatística no ligar a Matemática à realidade, proporcionando oportunidades para os alunos desenvolverem individualmente ou em grupo atividades interdisciplinares, sempre que possível com recurso às tecnologias disponíveis. Este documento enfatiza a importância da Estatística para permitir que os alunos melhorem a capacidade de avaliar afirmações de caráter estatístico, em particular, as que surgem nos diferentes meios de comunicação social, etc. Salienta também a contribuição do tema para o desenvolvimento do espírito crítico e de capacidades de comunicação dos alunos.

Na atual sociedade global prolifera a informação quantitativa (na publicidade, na comunicação social, etc.) e, em particular, a informação estatística está presente como forma de lhe dar credibilidade. As ideias estatísticas são também utilizadas em diversas áreas do saber. Todavia, vários estudos apresentam evidência de que, com frequência, os adultos não pensam estatisticamente para analisar e tomar decisões sobre questões pessoais importantes (Ben-Zvi & Garfield, 2004). Para esta situação contribuirá a formação estatística que se recebe na escola, uma vez que existe uma forte tendência para a valorização do cálculo e de aspetos técnicos, em detrimento de atividades como o planeamento de um estudo estatístico e a análise de dados e sua interpretação (e.g., Burrill, 2008; Cockcroft, 1982; Garfield & Ben-Zvi, 2008; Ponte & Fonseca, 2001; Ponte & Turkman, 2000; Scheaffer, 2000, 2006). As abordagens tradicionais ao ensino da Estatística não favorecem um desenvolvimento adequado da capacidade de raciocinar ou pensar estatisticamente (e.g., Ben-Zvi & Garfield, 2004; Burrill 2008; Pfannkuch & Wild, 2004). Scheaffer (2000), embora reconheça a utilidade dos procedimentos técnicos como ponto de partida no ensino da Estatística, refere que há necessidade de se ir além da sua mecanização, para uma genuína "compreensão de análises e comunicação dos resultados [estatísticos]" (p. 158), visando um patamar superior que corresponde à “reflexão” (p. 158). Alguns aspetos realçados por Burrill (2008) que devem ser tomados em conta no ensino e formação de professores na área de Estatística são: a sua especificidade nos diferentes tópicos, o uso de gráficos com vista à introspeção dos dados e o entendimento do papel da variação na compreensão da história de um conjunto de dados em análise.

Assim, a Estatística, com as suas noções e métodos, poderá ter uma dupla funcionalidade, cada uma igualmente importante na vida dos cidadãos: permitir a compreensão de uma variedade de aspetos da sociedade atual e facilitar a tomada de decisões no quotidiano de cada cidadão, onde a incerteza e variabilidade estão presentes (e.g., Almeida, 2000; Cobb & Moore, 1997; Franklin et al., 2007; Gattuso & Otaviani, 2011; Scheaffer, 2006).

Vários autores reconhecem que a Estatística não é um ramo da Matemática, uma vez que se trata de uma disciplina que possui as suas ferramentas e modos de pensamento próprios (e.g., Batanero, Diaz, Contreras & Roa, 2013; Batanero & Godino, 2005; Burrill 2008; Cobb & Moore, 1997; Scheaffer, 2006; Shaughnessy, 2006, 2007; Snee, 1993). Por exemplo, Shaughnessy (2006) menciona que, numa caraterização do pensamento matemático, usaria aspetos tais como: “procurar padrões, abstrair, generalizar, especializar, gerar e aplicar algoritmos” (p. 78), entre outros. Contudo, o autor considera que estes aspetos são, de um modo geral, diferentes dos cinco aspetos enumerados por Wild e Pfannkuch (1999) descritivos do pensamento estatístico, os quais incluem: (1) necessidade de dados; (2) conhecimento contextual sobre os dados; (3) atenção à variação; (4) ferramentas de modelação (históricas, estatísticas e probabilísticas); e (5) *transnumeração*[3]. Deste modo, o autor refere que quem ensina Estatística na escola deve estar consciente destes aspetos do pensamento específicos desta área relevantes para o ensino e aprendizagem ao nível do ensino básico e secundário.

[3] Expressão criada por Wild e Pfannkuch (1999) que segundo Shaughnessy (2006) foi a melhor expressão que os autores encontraram para exprimir um processo que “fosse além de uma mera transformação ou re-representação dos dados para aqueles casos onde características evidentes de um contexto são de repente descobertas” (p. 79).

Embora a atividade estatística tenha caraterísticas particulares, na perspetiva de Ponte e Fonseca (2001), esta é uma área importante da Educação Matemática, uma vez que tem uma "enorme expressão na atividade social" (p. 7), nomeadamente, nas ciências sociais e humanas. Além disso, trata-se de um tema escolar que se reveste de uma especificidade própria relativamente aos outros tópicos incluídos no currículo, dado que "o seu objeto não são conceitos simples como números ou figuras geométricas, mas agregados de objetos – amostras, coleções" (p. 7). Os autores consideram que os grandes objetivos do ensino da Estatística se enquadram nos objetivos do ensino da Matemática, "na ótica da sua utilização, em processos de investigação e em contextos de atividade social" (p. 7). A inclusão da Estatística na disciplina de Matemática proporciona mais vitalidade à mesma e pode favorecer o desenvolvimento de uma cidadania mais ativa dos alunos na atual sociedade (Gattuso & Ottaviani, 2011; Scheaffer, 2006).

A natureza abstrata dos conteúdos estudados em Estatística, tal como acontece em Matemática, e o facto dos temas estatísticos estarem dependentes de situações contextualizadas originam dificuldades de aprendizagem nos alunos e são obstáculos ao desenvolvimento do seu raciocínio estatístico (delMas, 2004). Rubin, Bruce e Tenney (1990) expõem a necessidade de se desenvolver no ensino da Estatística modos de apoiar "os alunos a manter o que pode interessar das suas intuições prévias e modificar o que está confuso para se fornecer uma base sólida, de forma a ser-se um sofisticado consumidor da informação estatística" (p. 319).

Sharma (2005) refere que os alunos do ensino secundário apresentam dificuldades em compreender questões relacionadas com o processo de amostragem, nomeadamente, na recolha e variabilidade de amostras. Os conceitos estatísticos de amostragem e variabilidade da amostra são complexos para os alunos (Rubin et al., 1990; Sharma, 2005). Além de recomendar um ensino que dê mais atenção a esses conceitos, Sharma (2005) insiste na necessidade de os professores darem mais atenção às respostas fornecidas pelos alunos, pois estes por vezes dão respostas corretas baseando-se em argumentos errados. Também Engel e Sedlmeier (2011) e Sorto, White e Lesser (2011) sugerem um ensino da regressão linear que permita que os alunos desenvolvam as suas intuições iniciais à volta da noção de reta de regressão para ideias mais sofisticadas que incluam a variabilidade dos dados e sua modelação.

Há vários estudos que indicam que os alunos podem aprender como calcular medidas formais de variabilidade, tais como, o desvio-padrão, os intervalos interquartis, etc., e porém continuarem a apresentar dificuldades de compreensão das suas diferentes representações e das relações com outros conceitos estatísticos (Garfield & Ben-Zvi, 2008).

O relatório *Matemática 2001* (APM, 1998) indica que o tema da Estatística e Probabilidades, do ensino secundário, foi o mais referenciado pelos professores de Matemática no sentido da sua exclusão ou simplificação no programa. O documento sugere, no âmbito da formação de professores, a necessidade de os sensibilizar para a importância dos temas que tinham vindo a ganhar maior ênfase nos programas da época, como era o caso da Estatística, devendo-se nomeadamente realçar "o seu papel formativo como instrumento na interpretação e intervenção sobre a realidade" (p. 30). É fundamental ter-se uma perceção das dificuldades e dos sucessos que professores e alunos experienciam no ensino e aprendizagem da Estatística de forma a progredir-se no seu ensino (e.g., Boaventura & Fernandes, 2004; Branco, 2000; Burrill, 2008; Fernandes, Carvalho & Correia, 2011; Shaughnessy, 2007).

Os estudos realizados por Groth e Bergner (2005, 2006), no âmbito da Estatística, sugerem que alguns professores evidenciam algumas dificuldades relativamente a conceitos estatísticos básicos, tais como, amostra e medidas de tendência central. Também Watson e Callingham (2013) encontraram dificuldades dos professores do ensino básico e do secundário em mobilizar eficientemente conhecimentos que possuíam sobre a média para ensinar. Batanero (2001) menciona também outros aspetos que dificultam a atividade do professor no ensino da Estatística, nomeadamente: ser uma ciência relativamente recente e em expansão; ser uma ciência de natureza interdisciplinar; existirem erros conceptuais em vários manuais escolares.

É amplamente reconhecido que a qualidade do ensino da Matemática não depende apenas do conhecimento matemático do professor (e.g., Ball, Hill & Bass, 2005; Batanero, 2001; Burgess, 2009; Ponte & Chapman, 2006; Shulman, 1986). Nesse sentido, Ball, Thames e Phelps (2005, 2008) mencionam a necessidade de os professores desenvolverem um conhecimento especialmente útil no ensino da Matemática: o conhecimento matemático para ensinar. Para caraterizar este conhecimento, estes autores consideram também imprescindível a observação das práticas do professor, de modo a tentar-se esclarecer e compreender o que é que os professores precisam de saber e ser capazes de fazer para ensinar Matemática. Inspirados no conhecido modelo de Shulman (1986) relativo ao conhecimento profissional do professor, estes autores incluem, no conhecimento matemático para ensinar, as seguintes quatro categorias: o conhecimento comum do conteúdo, o conhecimento especializado do conteúdo, o conhecimento do conteúdo e ensino, e o conhecimento do conteúdo e alunos.

No domínio do conhecimento profissional do professor há poucos quadros teóricos de referência na Estatística. Contudo, dada a emergência da Estatística na sociedade contemporânea e a necessidade da escola responder positivamente ao desafio de formar cidadãos que sejam, simultaneamente, consumidores e produtores de informação estatística (Almeida, 2000), torna-se pertinente o desenvolvimento de investigações neste domínio. Por conseguinte, há alguns investigadores que se têm dedicado a aprofundar e aperfeiçoar quadros teóricos usados na investigação em Educação Matemática, que passam a ter em conta elementos estatísticos. Tal é o caso de Groth (2007) que apresenta um modelo hipotético do conhecimento estatístico para ensinar e de Burgess (2006) que propõe um quadro teórico para perscrutar o conhecimento do professor "tal como é usado na prática letiva da Estatística" (p. 1). Estes modelos conceptuais têm em comum o facto de: (i) terem sido construídos com base nos trabalhos de Ball e colaboradores (2004, 2005) sobre o conhecimento matemático para ensinar, considerando alguns domínios deste conhecimento; e (ii) atenderem à especificidade do processo de investigação estatística. Numa publicação mais recentemente, Groth (2013) desenvolve o modelo anterior, integrando mais elementos do conhecimento matemático para ensinar de Ball e colaboradores (2005) e ainda mais dois construtos – *Compreensões chave do desenvolvimento* e *Ideias pedagogicamente poderosas* –, inspirando-se nos trabalhos de Simon (2006) e de Silverman e Thompson (2008), respetivamente. De acordo com Batanero (2000), é preciso experimentar e avaliar métodos de ensino adaptados à natureza específica da Estatística, "dado que nem sempre é possível transferir princípios gerais do ensino da Matemática" (p. 32).

Por sua vez, Batanero e Godino (2005) reconhecem que o conhecimento profissional do professor em Estatística deve incluir conhecimento estatístico mas também o que designam por conhecimento didático. Este conhecimento inspirado na noção de conhecimento pedagógico de conteúdo de Shulman (1986) engloba os seguintes aspetos complementares: (i) capacidade de reflexão epistemológica sobre o significado dos conceitos e procedimentos e sobre a natureza do conhecimento estatístico e sua evolução; (ii) capacidade de adaptação dos conteúdos nos diferentes níveis de ensino, o que contribuirá, por exemplo, para uma reflexão acerca dos diferentes níveis de compreensão dos alunos relativamente a um mesmo conceito; (iii) conhecimento dos interesses dos alunos, das dificuldades, erros e obstáculos nas aprendizagens e das suas estratégias na resolução de problemas, que permitirá uma melhor condução do ensino e avaliação da aprendizagem; (iv) capacidade de análise de diferentes tipos de recursos (currículo, situações didáticas, manuais, tecnologias) e uso de metodologias de ensino (tecnologias, tendo em conta a especificidade dos temas) que contribuam para melhorar as práticas. Esta descrição dos autores harmoniza-se com a noção de conhecimento didático de Ponte e Oliveira (2002) e de Canavarro (2003) (ver capítulo 3). No presente estudo adota-se a perspetiva destes autores sobre o conhecimento didático do professor. É assim importante realizar estudos que permitam compreender melhor os diferentes aspetos do conhecimento didático do professor de Matemática, bem como as conexões entre eles (Chapman, 2013), no domínio da Estatística.

De um modo muito breve, o modelo do *conhecimento didático em Estatística* é visto como estando ancorado na prática letiva, integrando a ação do professor e tendo em conta a especificidade da Estatística e do seu ensino (Batanero et al., 2013; Burgess, 2009; Shaughnessy, 2006; Wild & Pfannkuch, 1999). Trata-se de um modelo constituído por quatro domínios, *conhecimento de Estatística*, *conhecimento do currículo*, *conhecimento do aluno*, e *conhecimento de ensino do tema/tópico estatístico*, que se interrelacionam entre si. O *conhecimento de Estatística* diz respeito ao conhecimento que o professor tem dos aspetos específicos do saber que ensina (Godino et al., 2008), nomeadamente, dos conceitos estatísticos fundamentais, suas interligações, bem como ao conhecimento dos diversos dilemas e desafios envolvidos em lidar com as ideias estatísticas (Batanero et al., 2013; Burrill & Biehler, 2011). O *conhecimento do currículo* diz respeito ao conhecimento dos objetivos e finalidades do currículo escolar e sua articulação vertical e horizontal; este conhecimento tem um papel fundamental na tomada de decisões sobre os assuntos a que se deve dedicar mais tempo ou aprofundar e sobre a forma de orientar o processo de ensino-aprendizagem na unidade de Estatística (Godino et al., 2008; Ponte & Oliveira, 2002). O *conhecimento do aluno* diz respeito a uma compreensão sobre os interesses dos alunos, necessidades e dificuldades de aprendizagem frequentes (Godino et al., 2008; Ponte & Oliveira, 2002; Ponte, 2012); em particular, inclui a capacidade do professor prever erros e obstáculos enfrentados na aprendizagem (Godino et al., 2008; Groth, 2013), integra o conhecimento dos interesses dos alunos, estratégias, raciocínios, erros e dificuldades quando os alunos lidam com conceitos e com as tarefas propostas (Henriques & Ponte, 2014; Ponte & Oliveira, 2002; Ponte, 2012), e envolve a capacidade de compreender e avaliar o trabalho dos alunos (Burgess, 2008; Groth, 2013). O *conhecimento do ensino do tema/tópico estatístico* diz fundamentalmente respeito à condução efetiva das situações de ensino-aprendizagem (Ponte & Oliveira, 2002; Ponte, 2012), em particular ao modo como cada tópico estatístico é explorado em sala de aula e com que grau de exigência isso é feito (Godino et al., 2008); é um conhecimento que atende: às opções de tarefas, às interações que ocorrem em aula, à regulação geral do processo de ensino-aprendizagem e, ainda, às reflexões do professor sobre a sua prática (Godino et al., 2008; Henriques & Ponte, 2014; Ponte, 2012). Este último domínio ocupa o núcleo fundamental do modelo do *conhecimento didático em Estatística* visto ser essencialmente focado na prática, dada a sua centralidade e interações com as outras dimensões do modelo (Ponte, 2012) (ver detalhes no capítulo 3).

A literatura revela ainda um défice de estudos sobre o ensino da Estatística e, muito em particular, no ensino secundário (e.g., Batanero, Garfield, Ottaviani & Truran, 2001; Boaventura & Fernandes, 2004; Branco, 2000; Engel & Sedlmeier, 2011; Fernandes et al., 2011; Sharma, 2005). Contudo, há vários autores que enfatizam a relevância do desenvolvimento do raciocínio estatístico dos alunos em todos os níveis escolares, ou seja, desde o ensino básico ao secundário e mesmo no ensino superior (e.g., Ben-Zvi & Garfield, 2004; Burrill & Biehler, 2011; Shaughnessy, 2007). Fernandes (2009), para além de referir a importância do desenvolvimento de estudos sobre o ensino e aprendizagem da Estatística em Portugal, no ensino básico e secundário, alude a que, no contexto de formação e de desenvolvimento profissional do professor, ao nível da formação inicial, da formação contínua e da promoção de práticas colaborativas entre professores, há a necessidade de aprofundamento dos conteúdos estatísticos e da sua didática.

A melhoria da aprendizagem dos alunos passa também pelo desenvolvimento profissional dos professores. Para tal, é necessário criar oportunidades de formação através das quais os professores possam adquirir ou desenvolver conhecimento matemático e pedagógico relevante, num contexto em que estes desempenhem um papel ativo e que proporcione reflexão sobre as práticas (Sowder, 2007). Para Ponte (1998), o desenvolvimento profissional do professor ao longo da sua carreira é um aspeto que caracteriza indubitavelmente a profissão docente no mundo atual.

1.3 Problema e questões do estudo

Esta investigação insere-se na área do conhecimento profissional do professor de Matemática, mais especificamente no conhecimento didático, que se trata de um conhecimento orientado para situações da prática, perspetivado tendo em conta o pensamento e ação do professor (Canavarro, 2003; Ponte, 2008, 2012; Ponte & Oliveira, 2002). Este conhecimento no âmbito da Estatística inclui quatro domínios principais, nomeadamente, conhecimento de Estatística, conhecimento do ensino do tópico ou tema estatístico, conhecimento do aluno e conhecimento do currículo. O estudo tem como contexto o ensino da unidade de Estatística, na disciplina de Matemática A, no 10.º ano de escolaridade.

Na investigação optou-se pela escolha de professores que lecionassem no ensino secundário, pelo facto do tema da Estatística no programa desse nível (ME, 2001) abranger mais tópicos estatísticos e com maior grau de sofisticação do que o correspondente tema no ensino básico. Outro fator que foi ponderado na opção pela Matemática A, no 10.º ano, foi o facto de serem quase inexistentes estudos nesta disciplina e no ensino secundário, em particular a nível nacional.

O presente estudo incide especificamente no conhecimento didático do professor do ensino secundário e tem em conta vários aspetos deste conhecimento a partir da prática, relacionando conhecimento profissional e prática. O objetivo principal da investigação é compreender o conhecimento didático de duas professoras de Matemática relativamente aos diversos tópicos da unidade de Estatística na disciplina de Matemática A, no 10.º ano. Mais concretamente, com este estudo procura-se responder às seguintes questões:

1. Como se caracteriza o conhecimento didático em Estatística das professoras no ensino dos seguintes tópicos do programa:

- população e amostra;
- organização e interpretação de dados;
- medidas de localização;
- medidas de dispersão;
- dados bivariados e regressão linear?

2. Que aspetos transversais do conhecimento didático em Estatística das professoras se destacam na articulação entre os tópicos do programa?

1.4 Organização do estudo

Este trabalho está organizado em sete capítulos. O primeiro capítulo inclui a motivação, contexto e pertinência, problema e questões do estudo; o segundo e terceiro capítulos integram a fundamentação teórica do estudo; o quarto capítulo informa acerca das opções metodológicas adotadas na investigação; os dois capítulos seguintes apresentam os dois casos, professora Estela e professora Lia, respetivamente; o capítulo final, apresenta uma síntese do estudo, as principais conclusões do estudo e, adicionalmente, uma reflexão final, com comentários e implicações do estudo, bem como algumas sugestões para futuras investigações.

CAPÍTULO 2

O ENSINO E A APRENDIZAGEM DA ESTATÍSTICA

Este é um dos capítulos de enquadramento teórico do estudo e está organizado em três secções. A primeira, *Perspetivas sobre a estatística, o seu ensino e a sua aprendizagem*, inclui referências ao surgimento e importância da Estatística, nomeadamente, no contexto escolar; aos aspetos que aproximam e que afastam Matemática e Estatística e que devem ser tidos em conta no processo de ensino-aprendizagem; às definições mais usadas e reconhecidas na literatura acerca de literacia, raciocínio e pensamento estatístico. A segunda secção, *Ensino e aprendizagem de conceitos estatísticos*, apresenta uma revisão de literatura teórica e empírica focada em conceitos estatísticos centrais no ensino secundário, incluindo, em particular, referências a estudos que mostram algumas dificuldades e desafios no ensino e na aprendizagem destes conceitos. A terceira secção, *Orientações curriculares para a Estatística*, inclui fundamentalmente orientações curriculares nacionais para o ensino da Estatística no ensino secundário e também outras de relevo internacionais.

2.1 Perspetivas sobre a Estatística, o seu ensino e a sua aprendizagem

2.1.1 Surgimento e importância da Estatística

A palavra *Statistik* (Estatística) deve-se ao professor alemão e economista Gottfried Achenwall (1719-1792) que a introduziu, em 1749, e a relacionou com o uso de Matemática ao serviço da nação (Coolidge, 2013). De acordo com Ottaviani (1987), a palavra alemã *Statistik* provém da palavra *Status*, que em latim significa Estado. De modo semelhante, a palavra italiana *Statista* tem também um significado associado a Estado. A autora refere ainda que em Itália, a Estatística tem as suas origens numa disciplina intitulada *Ciência do Estado*[4] (Ottaviani, 2002). Nessa época, esta disciplina tinha como propósito principal "descrever assuntos de importância para um Estado ou Sociedade, de modo a fornecer ao Governo, e Administradores Públicos, o conhecimento e informação necessária para *bem governarem*" (Ottaviani, 2002, p. 30). Deste modo, segundo esta autora, é possível encontrar-se raízes do conhecimento estatístico com o começo da organização das sociedades; no entanto, a Estatística, enquanto disciplina, é relativamente recente (Branco, 2000). Por exemplo, Florence Nightingale[5], durante o século XIX, além de ter desenvolvido procedimentos de modo a que hospitais ingleses fossem capazes de dar informação estatística sobre os seus pacientes, apontou a Estatística como uma disciplina relevante a nível superior e também para estudar vários fenómenos sociais (Coolidge, 2013).

Desenvolvida sobretudo por matemáticos e cientistas de várias áreas (por exemplo, do mais recente para o mais antigo: Pearson, Galton, Quelelet, Gauss, LaPlace)[6], a Estatística teve grandes progressos do ponto de vista teórico e prático (no modo da sua utilização) no final do século XIX e, sobretudo, nos primeiros trinta anos do século XX (Branco, 2002; Coolidge, 2013). Surgiram, então, vários métodos estatísticos relevantes para a resolução de problemas das ciências aplicadas e, simultaneamente, assistiu-se "a uma generalizada emergência e reconhecimento de problemas de natureza estatística nos vários ramos científicos, na indústria, e em atividades governamentais, o que faz crescer o interesse pela atividade estatística" (Branco, 2000, p. 12). Neste contexto, surgiu a necessidade de se ensinar Estatística a um número crescente de pessoas, inicialmente através de cursos-treino para especialistas e utilizadores e, mais tarde, em cursos de pós-graduação e em cursos universitários (Branco, 2000).

[4] De acordo com a autora, esta disciplina foi lecionada, pela primeira vez, na Alemanha, em 1660, na Universidade de Helmstedt, pelo professor de Direito, F. Conring.

[5] Enfermeira e reformista de hospitais ingleses que contribuiu para o desenvolvimento da Estatística descritiva, foi também amiga de Francis Galton e contemporânea de Karl Pearson.

[6] *Karl Pearson*, matemático inglês, é reconhecido como o fundador da Ciência Moderna da Estatística, foi bastante influenciado pelo trabalho de Francis Galton; *Francis Galton*, cientista inglês, sobrinho de Charles Darwin, publicou uma obra nos finais do séc. XIX que espelha o início da era da Estatística Moderna; *Adolf Quetelet*, matemático belga e astrónomo, estendeu algumas ideias de Gauss para resolver problemas sociais, mais especificamente, para estudar o crime na sociedade*; Karl Gauss*, matemático alemão e astrónomo, contribuiu para o desenvolvimento da Estatística descritiva e inferencial; *LaPlace* (ou Pierre Simon), tal como era conhecido, matemático francês, foi pioneiro no desenvolvimento da Teoria das Probabilidades e, consequentemente, das bases da Estatística inferencial.

Já o interesse por questões de ensino da Estatística começa a aparecer no início do século XX mas lentamente (Branco, 2000). Este autor refere que no encontro anual da *American Statistical Association* (ASA), em 1925, esta problemática foi abordada em apenas duas sessões. Contudo, só em 1948 foi criada a comissão para a Educação do *International Statistical Institute* (ISI), com o propósito de desenvolver a Educação Estatística, em particular, promovendo a formação nesta área. Com este desígnio, esta comissão para a Educação do ISI colaborou também com a UNESCO e com outros organismos internacionais para desenvolver a Educação Estatística à escala mundial. Deste modo, houve logo por parte desta comissão a preocupação de se criar centros internacionais de Educação Estatística (nomeadamente, em 1950, o de Calcutá, na Índia) e na preparação de materiais estatísticos, tais como livros de texto universitários e dicionários de termos estatísticos, entre outros. O centro internacional de Estatística de Calcutá tinha como finalidade fornecer cursos de Estatística teórica e aplicada a participantes de vários países africanos e asiáticos (Batanero, 2001; Branco, 2000).

De acordo com Branco (2000), decorreu uma conferência em Royaumont, em França, nos finais de 1959, preparada pela Organização Europeia de Cooperação Económica (OECE, sucedida pela Organização de Cooperação e Desenvolvimento Económico, OCDE, em 1961), em que participaram vários matemáticos de todo o mundo. Embora com a finalidade "de estudar uma reforma profunda do ensino da Matemática ao nível do ensino pré-universitário" (p. 14), teve um papel fundamental para o avanço do ensino da Estatística. Nesta conferência concluiu-se, em particular, que os estudos secundários deveriam incluir um estudo elementar do cálculo das Probabilidades e de Estatística. Esta conclusão foi sustentada não só pela importância da Estatística na resolução de problemas, mas "também na necessidade de dar ao homem comum uma preparação adequada para que ele não seja surpreendido com um conjunto de noções básicas de Estatística que passaram a fazer parte da vida diária" (Branco, 2000, p. 15).

Com base nas orientações e recomendações de Royaumont foi publicado pela OCDE, em 1961, o livro intitulado *Um programa moderno de Matemática para o ensino secundário* que inclui como terceiro grande tema a Estatística e Probabilidades, seguindo-se à Álgebra e à Geometria. Este programa incorpora na Estatística, em termos de conteúdos, sobretudo noções básicas da Estatística descritiva e o estudo da Estatística inferencial (Guimarães, 2003, referindo OECE, 1961b). De acordo com Guimarães (2003), a conferência de Royaumont

> é certamente a realização mais emblemática de todo o movimento reformador de grande influência internacional que recebeu o nome de Matemática Moderna e, também, uma das mais conhecidas na história da revolução curricular recente do ensino da Matemática. (p. 91)

O movimento que se começou a registar em diversos países da Europa, e também noutros continentes, com o objetivo de reformar os programas do ensino secundário ao nível dos conteúdos matemáticos e dos métodos de ensino, chegou a Portugal em meados dos anos sessenta. O matemático português José Sebastião e Silva ficou com a responsabilidade de desenvolver este projeto (Guimarães, 2003, 2010). Em 1963/64 funcionaram, inicialmente, três turmas-piloto, a título experimental, no 3.º ciclo dos liceus, atuais 10.º e 11.º anos (Branco, 2000); este número foi-se alargando, nos anos seguintes, a várias escolas por todo o país (Branco 2000; Guimarães, 2010). Esta iniciativa, que se desenvolveu de modo gradual, permitiu a inclusão definitiva de novos tópicos no currículo de Matemática do ensino secundário, nomeadamente, Elementos de Cálculo, Probabilidades e Estatística (Branco, 2000).

Contudo, para Ponte e Turkman (2000), o facto de não ter sido providenciado, desde logo, um apoio adequado aos professores para o ensino dos novos temas de Estatística e Probabilidades do currículo, "permitiu que estes tópicos fossem os primeiros a ser sacrificados quando é necessário fazer cortes no programa" (p. 5). Apesar disto, as inúmeras ações de formação promovidas para professores nesta área, o continuado reforço destes tópicos nos currículos e uma maior tomada de consciência pelos professores do papel da Estatística no desenvolvimento global dos alunos são aspetos que foram promovendo alguma mudança (Ponte & Turkman, 2000).

Segundo Batanero (2001), na maioria dos países desenvolvidos, a Estatística começa a ser incorporada nos currículos de Matemática no ensino primário estendendo-se ao ensino secundário e a alguns cursos universitários. Por conseguinte, começam a aparecer projetos curriculares com o propósito de fomentar a educação estatística, por exemplo, no Reino Unido, o *School Council Project on Statistical Education* (1957-1981) e, nos Estados Unidos da América, o *Quantitative Literacy Project* (1985-1998) e o *Data Driven Mathematics* (1996-2000).

O ISI tem promovido conferências sobre a Educação Estatística, dando origem à primeira *International Conference on Teaching Statistics* (ICOTS), em 1982, e que se realiza desde então de quatro em quatro anos. Um acontecimento importante para o desenvolvimento da Educação Estatística foi a criação, pelo ISI, em 1991, de uma nova secção, designada por *International Association of Statistical Education* (IASE) que tem como propósito principal contribuir para o avanço da Educação Estatística a nível internacional. Atualmente, esta secção continua a promover diversos encontros e tem várias publicações importantes na área.

Na década de 90, intensificou-se a investigação na área da Educação Estatística estimulada pelos encontros e congressos do ISI, conferências do *International Congress on Mathematical Education* (ICME), das *Round Table Conferences* associadas ao ICME, de revistas tais como *Teaching Statistics* e *Journal of Statistical Education*, e de toda a informação relacionada com esta área disponibilizada na internet (Branco, 2000; Batanero, 2001). De acordo com Snee (1993), no início dos anos 90, em conferências promovidas pela *American Statistical Association* (ASA) e em vários artigos publicados nas revistas *The American Statistician* e *Amstat News*, surgem depoimentos de que a Educação Estatística não está no bom caminho, havendo a necessidade de uma mudança profunda nesta área. De facto, segundo este autor, de "um modo geral, as pessoas não compreendem o pensamento estatístico e, consequentemente, não dão valor ao seu uso" (p. 149). Refere ainda ser cada vez mais consensual que os conteúdos estatísticos se devem centrar mais na recolha de dados, na compreensão e modelação da variação estatística, na apresentação gráfica dos dados, e na resolução de problemas, entre outros aspetos, em vez de se focarem em aspetos matemáticos e probabilísticos.

Também Cobb e Moore (1997) destacam que no ensino da Estatística se deveria dar um maior realce aos dados e a conceitos estatísticos. Além disso, referem a importância do desenvolvimento dos aspetos básicos do pensamento estatístico nos alunos, sendo eles: a necessidade de dados, a importância da produção de dados, a omnipresença da variabilidade, a quantificação e explicação da variabilidade.

Snee (1993) sustenta ainda que, ao estabelecer-se maior ligação com problemas da vida real, os alunos dão maior sentido às ferramentas e aos conceitos estatísticos, e ficam mais próximos de experienciar o pensamento estatístico. O autor menciona não haver tanto consenso relativamente ao modo como o ensino da Estatística deve ser preconizado na prática. Contudo, considera essenciais as ideias de alguns investigadores na área, tais como Joan Garfield e George W. Cobb, ao considerarem que os alunos devem ter um papel ativo na aprendizagem da Estatística, “experienciando a Estatística” (Snee, 1993, p. 150). Esta experiência passa, por exemplo, pelo trabalho com dados reais e pela resolução de problemas reais. Cobb e Moore (1997) consideram que a análise exploratória de dados é fundamental sobretudo com dados reais:

> A maioria dos dados reais contém surpresas, algumas das quais podem invalidar ou forçar a modificação da inferência planeada. Isto é uma razão porque correr os dados, procedimento inferencial sofisticado, antes de se explorar os dados de modo cuidado é a marca de um noviço estatístico. (p. 848)

Por conseguinte, sem se proceder à exploração dos dados antes de uma análise mais formal, pode-se perder informação crucial, nomeadamente, a que poderia indicar qual ou quais os procedimentos analíticos mais apropriados para os dados em estudo. Adicionalmente, Moore (1998) destaca que no ensino da Estatística os alunos devem experienciar investigações estatísticas para que o seu raciocínio estatístico se desenvolva. De facto, para Gattuso e Ottaviani (2011), o conhecimento das teorias estatísticas não é suficiente para um ensino efetivo da Estatística. O professor deve ter oportunidade para desenvolver o seu pensamento estatístico, o que passa por experienciar a análise de dados e compreender os conceitos chave do pensamento estatístico, tal como é o caso da variabilidade. Deste modo, o professor deve possuir um conhecimento mais alargado e profundo do que aquele que vai ser ensinado aos alunos. Por conseguinte, atualmente há um aumento de investigações no ensino da Estatística com foco na literacia, no raciocínio, no pensamento estatístico e no desenvolvimento da compreensão concetual em Estatística (Garfield & Ben-Zvi, 2007).

2.1.2 A Matemática e a Estatística: proximidade e distância

De acordo com Cobb e Moore (1997), a Estatística tem um caráter metodológico na medida em que a sua existência se justifica com o fim de “oferecer aos outros campos de estudo um conjunto de ideias coerentes e ferramentas para se lidar com os dados. A necessidade de tal disciplina surge da omnipresença da variabilidade” (p. 801). A variabilidade está presente na realidade de diversas formas, por exemplo, nas medições repetidas a uma mesma característica de um dado objeto, e a Estatística está apetrechada de meios para lidar com a variabilidade presente nos dados. Este é um dos aspetos mais importantes que distingue o pensamento estatístico do matemático (Cobb & Moore, 1997).

A variabilidade que, de uma maneira geral, pode ser designada pela "propensão das observações de um conjunto de dados mudar" (Gattuso & Ottaviani, 2011, p. 127), existe na Matemática e na Estatística. Porém, é um conceito que tem um entendimento diferente nessas duas disciplinas. Por exemplo, numa aula de Matemática os alunos estudam a dependência de uma variável sobre outra e tentam traduzi-la matematicamente. Na Estatística, para se chegar à compreensão de uma distribuição, não basta ter em conta medidas de localização, a variabilidade dos dados tem de ser obrigatoriamente contemplada na análise (Gattuso & Ottaviani, 2011).

Para Rossman, Chance e Medina (2006), a Estatística é uma ciência que pretende obter um conhecimento profundo dos dados e que faz uso de muita Matemática. Contudo, à semelhança de Cobb e Moore (1997), estes autores consideram que a Estatística não é um ramo da Matemática e, por isso, apresentam cinco aspetos da Estatística que permitem evidenciar as diferenças entre estas duas disciplinas; são eles: i) o papel crucial do contexto; ii) questões de medição; iii) a importância da recolha de dados; iv) a falta de conclusões definitivas e; v) a comunicação do conhecimento estatístico.

Vejamos como podem ser interpretados cada um destes aspetos específicos da Estatística. No que diz respeito ao *primeiro aspeto* (o papel crucial do contexto), há a considerar que, enquanto a Matemática é uma ciência que pode ser estudada independentemente do contexto, o mesmo não se passa com a Estatística, onde este tem um papel fundamental na análise dos dados. Por exemplo, na Matemática, as funções lineares podem ser estudadas com base nas suas propriedades matemáticas sem se recorrer às aplicações; na Estatística, não se pode interpretar ou tirar qualquer conclusão sobre um simples gráfico de pontos sem qualquer referência contextual. Por exemplo, num estudo em que se pretendia averiguar a existência de relação entre a idade com que as crianças falaram pela primeira vez e o seu respetivo resultado num teste de inteligência (*Gesell aptitude test*) feito durante a infância, o diagrama de dispersão (figura 1) com a sua reta de regressão revela uma associação negativa entre as variáveis idade, em meses (eixo dos *xx*) e a pontuação obtida no teste (eixo dos *yy*). Tal poderia parecer indicar que uma idade mais tardia tenderia a estar associada a resultados mais baixos.

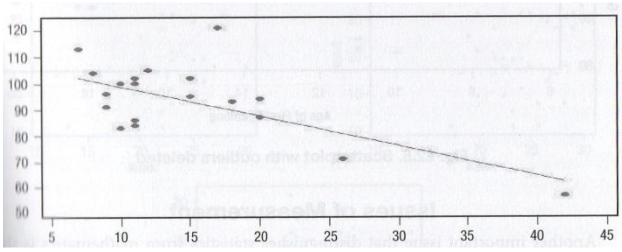


Figura 1: Diagrama de dispersão dos dados idade *versus* nota de teste.

No entanto, o contexto é primordial na decisão se os dados correspondentes às duas crianças que começaram a falar tardiamente e tiveram os resultados mais baixos no teste de inteligência serão *outliers* da distribuição. Ao considerar-se a eliminação desses dois dados e traçar-se novamente a reta de regressão sobre o conjunto de dados, conclui-se que não há associação entre as duas variáveis (figura 2).

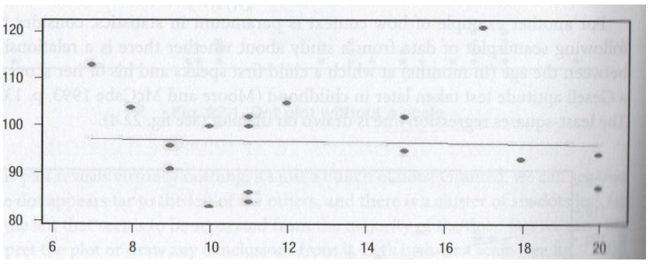


Figura 2: Diagrama de dispersão sem os *outliers* e com nova reta de regressão.

Cobb e Moore (1997) consideram que o contexto na Matemática "obscurece a estrutura" (p. 803), enquanto na análise estatística de dados "proporciona significado" (idem). Adicionalmente, estes autores referem que na Matemática os professores habitualmente inventam exemplos contextuais que parecem resultar; na Estatística exemplos desta natureza não têm o mesmo sucesso pois "não fornecem uma genuína interação entre padrão e contexto" (idem). Na Estatística, o contexto fornece o significado através do qual os padrões (nos dados) podem ser examinados e talvez explicados. O trabalho na

Estatística exige que se conheça a natureza dos dados, como e onde é que estes dados foram produzidos, de modo a prosseguir-se com a análise e chegar-se a conclusões. Já na Matemática o contexto pode ser usado, por exemplo, com o intuito de trazer motivação para a sala de aula (Gattuso & Ottaviani, 2011).

Quanto ao *segundo aspeto* mencionado por Rossman et al. (2006), a medição em Matemática tem um foco diferente da medição em Estatística. Na primeira, a medição envolve levar os alunos a aprender unidades apropriadas para medir características de um objeto tais como comprimento, área ou volume e usar fórmulas para medir estas caraterísticas. Ainda na Matemática, por exemplo na geometria, para mostrar que dois lados de um triângulo têm o mesmo comprimento, pode não haver necessidade do uso de uma régua, dado que essa igualdade pode, de facto, ser "deduzida através de hipóteses, definições e teoremas" (Gattuso & Ottaviani, 2011, p. 127). Já na Estatística, tirar conclusões dos dados depende, de uma forma decisiva, de fazer medições válidas das propriedades a serem estudadas. Como se trata de uma disciplina que é "sobretudo sobre compreensão, medição e descrição do mundo real, tirar medições válidas é crucial" (Gattuso & Ottaviani, 2011, p. 127). Em Estatística também há questões de medição que surgem quando o foco de um estudo estatístico é, por exemplo, o desemprego, capacidade de memória, eficácia do ensino, etc., uma vez que nem sempre é fácil saber-se ou detetar-se quais os aspetos que são realmente relevantes medir em cada uma destas situações.

O *terceiro aspeto* referido por Rossman et al. (2006), e que distingue a Matemática da Estatística, prende-se com a recolha de dados. Mesmo que em Matemática se analise um conjunto de dados, o propósito principal desta tarefa será encontrar e analisar algum padrão deste conjunto de dados sendo completamente desnecessária qualquer informação adicional sobre o modo como estes dados foram recolhidos. Em Estatística, esta mesma informação é fundamental para se saber que tipo de resultados se pode aspirar obter. Por exemplo, conclusões de causa e efeito não podem ser obtidas de um estudo que não tenha por base uma experiência aleatória.

Relativamente ao *quarto aspeto* – falta de conclusões definitivas em Estatística – Rossman et al. (2006) consideram que a "A Estatística e a Matemática perguntam diferentes tipos de questões e, consequentemente, alcançam conclusões de natureza diferente" (p. 329). Enquanto a Matemática envolve o uso do raciocínio dedutivo rigoroso e, por conseguinte, chega-se a soluções corretas e sucintas para questões às quais se pretende dar resposta, na Estatística, por seu turno, usa-se o raciocínio indutivo e, consequentemente, obtêm-se conclusões, com um certo grau de incerteza. De facto, nas conclusões estatísticas lêem-se mais frequentemente frases do tipo *os dados sugerem fortemente que* em vez de *os dados provam que*.

Além disso, é importante que haja consciencialização de que em Estatística podem existir diversas abordagens para analisar os dados, podendo estas, por vezes, originar conclusões diferentes à luz das questões e do contexto em que os dados foram recolhidos. A mesma questão com os mesmos dados pode conduzir a análises e soluções diferentes; contudo, há sempre a necessidade de se explicar e defender as conclusões obtidas (Gattuso & Ottaviani, 2011; Rossman et al. 2006).

Finalmente, quanto ao *quinto aspeto*, relativo à comunicação do conhecimento estatístico, Rossman et al. (2006) mencionam que a comunicação é importante quer para a Matemática quer para a Estatística. Contudo, em Matemática esta é frequentemente realizada em forma simbólica, enquanto em Estatística o problema a ser estudado deve estar formulado ou entendido de modo claro e, depois de realizada a análise dos dados, devem os resultados ser comunicados de modo percetível. Tal como acontece em Matemática, em Estatística há muitos termos estatísticos que são usados pelos alunos na linguagem corrente; daí estes transportarem, frequentemente, os significados desses termos para a Estatística. Só com muita prática e reflexão sobre estas questões é que os alunos podem reaprender os novos significados técnicos destes termos, de acordo com a linguagem estatística, e ultrapassar essas dificuldades de terminologia. Esta aprendizagem influencia positivamente a comunicação do conhecimento estatístico (delMas, 2004; Rossmam et al., 2006; Shaughnessy, 2007).

Ainda assim, de acordo com Scheaffer (2006), é possível fazer-se um casamento feliz entre a Estatística e a Matemática, embora o autor reconheça que estas duas disciplinas sejam distintas uma da outra. De facto, este autor considera que o pensamento matemático é distinto do pensamento estatístico. O primeiro é um pensamento que é frequentemente determinista, na medida em que "todo o resultado tem de ter uma causa explicável" (p. 310), enquanto o segundo segue um pensamento frequentemente probabilístico no qual “um resultado pode ser fruto de um de vários fatores inexplicáveis que se conjugam, cujo efeito resultante se chama acaso” (p. 310).

Para Scheaffer (2006), a Educação Estatística nas escolas continua a ser um assunto importante e atual, e são vários os documentos que o comprovam, nomeadamente os *Principles and standards for school mathematics* (NCTM, 2000) e, mais recentemente, o relatório *Ready or not: Creating a high school diploma that counts* (American Diploma Project, 2004). Gattuso e Ottaviani (2011) e Scheaffer (2006) consideram que o modo como a Estatística é ensinada nas escolas pode conduzir os alunos a uma visão determinista do mundo quantitativo que rodeia os indivíduos.

Assim, para Scheaffer (2006) é crucial que na escola seja dada a devida importância à Estatística, nas aulas de Matemática, de modo a que os alunos aprendam a diferenciar entre o pensamento matemático e o estatístico. Os materiais curriculares muitas vezes ofuscam essas diferenças, contribuindo para que a confusão permaneça. Este autor apresenta alguns exemplos de tarefas que são habitualmente propostas aos alunos nas aulas de Estatística, umas de foco matemático e outras de foco estatístico, que se passam a apresentar de seguida.

> **Exemplo 1:** Os alunos da aula do professor António votaram no livro favorito que leram para esta aula, durante os últimos três meses. Os livros lidos foram: A, B e C. Algumas informações úteis sobre a votação obtida são as seguintes:
>
> - 34 alunos votaram;
> - O livro vencedor obteve a maioria dos votos, mas obteve um pouco menos do que metade dos votos;
> - Houve empate para o segundo lugar.
>
> Faça uma tabela ou um gráfico que mostre os resultados dos votos que satisfaçam estas três informações dadas.
>
> (Adaptado de Scheaffer, 2006, p. 316)

> **Exemplo 2:** O André obteve os seguintes resultados em três testes, 78, 76 e 74, enquanto a Maria obteve 72, 82 e 74. Compare a média dos resultados destes dois alunos e selecione a resposta apropriada:
>
> - A média do André é um ponto mais alta que a da Maria;

- A média do André é um ponto mais baixa que a da Maria;
- As médias do André e da Maria são iguais;
- A média do André é dois pontos mais alta que a da Maria;
- A média do André é dois pontos mais baixa que a da Maria.

(Adaptado de Scheaffer, 2006, p. 317)

Para Scheaffer (2006), o primeiro exemplo não é uma questão sobre análise de dados, é uma questão sobre propriedades de números, ou seja, que envolve um raciocínio matemático em vez de um raciocínio estatístico. Para o autor, o raciocínio estatístico, por sua vez, compreende o "uso dos números no contexto – dados – para responder a uma questão prática de interesse" (p. 316). Assim, este autor considera que o segundo exemplo, embora diga respeito ao cálculo de duas médias e sua comparação, não é uma questão de análise de dados no sentido do raciocínio estatístico, uma vez que o contexto da questão é bastante superficial e dispensável.

O autor apresenta mais dois exemplos, o exemplo 3 e o exemplo 4. Com o exemplo 3 pretende realçar uma tarefa muitas vezes tratada sobretudo com foco nos procedimentos, dando uma indicação de como esta tarefa poderia ser usada no sentido de promover o raciocínio estatístico dos alunos. Com o exemplo 4 destaca uma situação que tem também potencial para promover o raciocínio e pensamento estatístico dos alunos, uma vez que são fornecidos dados reais, que devem ser analisados com as ferramentas estatísticas que os alunos possuem, cuja variabilidade não pode ser descurada, sendo igualmente necessário dar-se resposta à questão colocada.

Exemplo 3: A tabela abaixo mostra o número de pessoas que foram aos cinemas A e B durante 5 dias consecutivos e inclui a média e a mediana destas presenças durante estes 5 dias em cada um dos cinemas.

	A	**B**
Dia 1	100	72
Dia 2	87	92
Dia 3	90	70
Dia 4	10	71
Dia 5	91	100
Média	75.6	82
Mediana	90	72

Qual a estatística (média ou mediana) que deverá ser usada para descrever um dia típico de presenças para os cinco dias no teatro A? Justifique a sua resposta. Qual a estatística (média ou mediana) que deverá ser usada para descrever um dia típico de presenças para os cinco dias no teatro B? Justifique a sua resposta.

(Adaptado de Scheaffer, 2006, p. 318)

Exemplo 4: A tabela seguinte indica os tempos registados em sete corridas de 100 metros, deste ano, de três atletas do sexo feminino. Apenas uma delas pode entrar numa próxima competição. Qual das atletas escolheria para participar na próxima competição? Justifique a sua resposta.

	1	**2**	**3**	**4**	**5**	**6**	**7**
Susana	15.2	14.8	15.0	14.7	14.3	14.5	14.5
Tânia	15.8	15.7	15.4	15.0	14.8	14.6	14.5
Lara	15.6	15.5	14.8	15.1	14.5	14.7	14.5

(Adaptado de Scheaffer, 2006, p. 319)

O exemplo 3 refere-se novamente a uma questão algorítmica (tal como a do exemplo dois) que foi elaborada para os alunos responderem quase de modo automático e pensarem do seguinte modo: sempre que um conjunto de dados parecer conter um *outlier*, a mediana deve ser a medida escolhida para descrever a situação, dado ser a medida entre as duas estatísticas referidas como mais resistente a *outliers*; caso contrário, é conveniente escolher-se a média como estatística sumária (Scheaffer, 2006). Contudo, este exemplo daria uma boa questão estatística se mais informação sobre o contexto do problema fosse providenciada. Como refere o autor:

> Medidas estatísticas apropriadas não podem ser determinadas com estes dados até o contexto ser constituído por informação suficiente para se compreender como e porquê os dados foram recolhidos e qual a questão que precisa de ser respondida com estes dados. (p. 318)

O exemplo 4 é uma questão aberta que vai permitir aos alunos explorarem os dados realistas fornecidos com base nas ferramentas estatísticas que possuem, de modo a responder à questão colocada e, após esta análise de dados, comunicarem os resultados de modo coerente. Portanto, os alunos terão de dar resposta à questão usando aspetos do pensamento estatístico.

Ainda de acordo com Scheaffer (2006), há grandes vantagens se a Estatística e a Matemática forem lecionadas no âmbito da disciplina de Matemática. Deveras, a Estatística é reconhecida pelo seu papel na tomada de decisões quer pessoais quer profissionais na sociedade moderna, enquanto a Matemática é uma disciplina com história que ocupa uma posição importante no currículo escolar. Para além disso, acrescenta:

> A Matemática e a Estatística não se deveriam distanciar uma da outra. A Estatística seria então espalhada entre diversas áreas de aplicação que necessitam apenas de algumas partes da disciplina e ficaria enfraquecida como disciplina unificada com a sua integridade disciplinar. A Matemática perderia a vitalidade de ver os seus métodos usados numa importante e interessante área de aplicação. (p. 314)

Segundo Rossman et al. (2006), aqueles cinco aspetos referidos anteriormente (o papel crucial do contexto; questões de medição; a importância da recolha de dados; a falta de conclusões definitivas e a comunicação do conhecimento estatístico) podem ajudar a compreender algumas diferenças entre a Estatística e Matemática, apesar de não serem aspetos necessariamente universais. Estes podem emergir nas aulas em que se ensina conteúdo matemático, mas é naquelas onde se ensina a Estatística que eles devem ser efetivamente centrais. Os professores devem estar a par das dificuldades dos alunos na Estatística, ajudá-los e orientá-los na construção deste conhecimento, envolvendo-os em atividades que fomentem o espírito crítico, a comunicação e o raciocínio estatístico. Adicionalmente, apesar de Rossman et al. (2006) terem plena consciência que há muito a fazer pelo conhecimento e desenvolvimento profissional do professor em Estatística, mencionando o relatório *A educação matemática dos professores* (da *Conference Board of the Mathematicals Sciences*, 2001), indicam duas razões principais para a necessidade de os professores se preocuparem com o ensino da Estatística: “um diferente tipo de preparação ao nível da instrução é necessário para o ensino da Estatística; os alunos reagem de forma diferente à Estatística e à Matemática” (p. 331).

A Estatística permite que os alunos desenvolvam certas capacidades de resolução de problemas, nomeadamente, colocar questões, analisar, representar e comunicar informação quantitativa que são também capacidades relevantes para a Matemática. Deste modo, Gattuso e Ottaviani (2011) consideram que os professores devem ter simultaneamente consciência dos benefícios do ensino da Estatística na disciplina de Matemática e da especificidade de cada disciplina. Estas autoras também declaram que para os alunos que não estejam muito motivados para a Matemática, os "tópicos estatísticos oferecem uma oportunidade de começar de fresco e de abrir os olhos dos alunos para uma nova perspetiva nas ciências matemáticas" (p. 333), sobretudo quando esses tópicos são trabalhados com base numa situação real, que tenha em conta os interesses dos alunos, situação que pode criar motivação e maior envolvimento na aula de Matemática.

Para Ponte e Fonseca (2001) e Ben-Zvi e Garfield (2004), na área da Estatística, ao nível do ensino básico e secundário, existem orientações curriculares americanas e inglesas que colocam em primeiro plano a análise de dados. Ben-Zvi e Garfield (2004) consideram que vários fatores têm influenciado o desenvolvimento do ensino da Estatística em todos os níveis escolares, nomeadamente:

> (1) As mudanças no campo da Estatística, incluindo as novas técnicas de exploração de dados, e a crescente disponibilidade tecnológica nas escolas e em casa;
> (2) Uma maior perceção das dificuldades dos alunos ao nível do raciocínio e pensamento estatístico, apesar dos bons resultados obtidos na Estatística;
> (3) Uma maior preocupação com a formação de professores, de todos os níveis de escolaridade, que lecionam Estatística (inclusive ao nível universitário) devido ao facto de muitos nunca terem estudado Estatística nem terem experienciado atividades de análise de dados. (p. 5)

No contexto português, o relatório *Matemática 2001* (APM, 1998) indicou que o tema da Estatística e Probabilidades, do ensino secundário, foi o mais apontado pelos professores no sentido da sua exclusão ou simplificação do programa. Também um estudo exploratório, mais recente, sobre perspetivas e práticas no domínio do ensino da Estatística no secundário, que envolveu a aplicação de um questionário dirigido a professores (*n*=110), cobrindo cerca de 90% das escolas secundárias afetas à Direção Regional de Educação do Norte, revela que um número significativo destes professores (cerca de 40%) atribui uma importância de *reduzida a média* à Estatística na formação dos alunos, assim como à sua natureza muito particular no campo da Matemática (Quintas, Oliveira & Ferreira, 2009). A maioria dos professores inquiridos neste estudo considera que a Estatística do 10.º ano é um tema em que se sentem à vontade ensinar e em que, habitualmente, os alunos têm sucesso. Apenas cerca de 27% dos professores afirmam ter frequentado alguma formação na área da Estatística no período de 2000 a 2009, no entanto, cerca de 67% indica necessitar de formação nessa área, sobretudo ao nível da pedagogia. Quando se pronunciaram sobre a adequação do peso relativo do tema Estatística face aos restantes temas lecionados na Matemática A do 10.º ano, cerca de 65% dos professores afirma que o peso do tema Estatística, no 10.º ano, devia manter-se; 26% refere que o tema devia ser mais desenvolvido e cerca de 5% indica que devia ser excluído. Na aprendizagem da Estatística, a maioria dos professores valoriza a organização, representação e tratamento de dados em detrimento do enfoque na aquisição de técnicas de cálculo pelos alunos. Os professores também privilegiam na aula as atividades de resolução de problemas e de exercícios em desfavor de atividades de exploração e trabalho de projeto. Os recursos mais utilizados nas suas aulas de Estatística são o manual adotado, fichas de trabalho produzidas pelo professor e a calculadora gráfica; pelo contrário, a internet e a folha de cálculo têm uma frequência de utilização baixa na aula (Quintas et al., 2009).

Estas autoras acrescentam ainda que os dados quantitativos deste estudo sugerem como pertinente a ideia de se continuar a reforçar o conhecimento do professor no ensino da Estatística, no ensino secundário, com vista ao desenvolvimento do raciocínio estatístico dos alunos. A Estatística tem de ser encarada como um tema curricular, da disciplina de Matemática, fundamental na formação dos cidadãos, dado poder capacitá-los para a compreensão e interpretação da informação estatística, assim como para serem interventivos no mundo que os rodeia (Ponte & Fonseca, 2001).

2.1.3 Literacia, raciocínio e pensamento estatístico

Segundo Ben-Zvi e Garfield (2004), ainda não existem definições consistentes de literacia, raciocínio e pensamento estatístico. Muitos autores substituem, facilmente, literacia estatística por literacia quantitativa e utilizam o raciocínio estatístico e o pensamento estatístico para designarem as mesmas capacidades cognitivas ou os mesmos tipos de atividade mental. De acordo com Ben-Zvi e Garfield (2004), estas questões de terminologia foram notórias, em 1998, na *Fifth International Conference on Teaching Statistics* (ICOTS-5), realizada em Singapura. Para os autores, há a necessidade de tentar chegar a definições consensuais para contribuir, efetivamente, para o desenvolvimento do ensino da Estatística.

Em 1999, foi organizada uma pequena conferência, *The First International Research Forum on Statistical Reasoning, Thinking and Literacy* (SRTL-1), em Israel, com a finalidade de se estudar as semelhanças e diferenças entre literacia, raciocínio e pensamento estatístico. O fórum seguinte (SRTL-2) realizou-se na Austrália, em 2001, e focou-se nos diferentes tipos de raciocínio estatístico. Em torno desta problemática, delMas (2002), refere o seguinte:

> Enquanto estas áreas [literacia, raciocínio e pensamento estatístico] representam objetivos de aprendizagem importantes para a instrução estatística, a falta de transparência na definição pode resultar numa falta de ligação entre o que ensinamos, o que os alunos aprendem, e o que avaliamos. (p. 1)

Na perspetiva de delMas (2004), os três processos cognitivos, literacia, raciocínio e pensamento estatístico, podem estar envolvidos em qualquer conteúdo estatístico; contudo, é possível que, na realização de uma tarefa, estes processos possam ser distinguidos pela natureza da própria tarefa. E exemplifica do seguinte modo:

> Uma pessoa que sabe quando e como aplicar o conhecimento e procedimentos estatísticos demonstra pensamento estatístico. Contrariamente, uma pessoa que saiba explicar como os resultados foram produzidos e porque é que uma conclusão está justificada demonstra raciocínio estatístico. (p. 85)

DelMas (2002) apresentou um quadro (quadro 1) no qual inclui uma lista de palavras que podem dar indicações sobre o propósito das tarefas propostas (ou questões colocadas) aos alunos. Além disso, tais indicações podem contribuir para se distinguir se as capacidades que estão a ser promovidas na realização destas tarefas estão mais ao nível da literacia, do raciocínio ou do pensamento estatístico ou se, eventualmente, mais do que um desses processos está a ser fomentado em simultâneo.

Quadro 1: Lista de palavras para distinguir os processos cognitivos envolvidos numa tarefa (adaptado de delMas, 2007, p.7)

Literacia estatística (básica)	**Raciocínio estatístico**	**Pensamento estatístico**
Identificar Descrever Parafrasear Traduzir Interpretar Ler	Porquê? Como? Explicar um processo	Aplicar Criticar Avaliar Generalizar

Por conseguinte, de acordo com delMas (2002, 2004), se o propósito do ensino e aprendizagem da Estatística é desenvolver nos alunos a *literacia estatística*, devem-lhes ser propostas questões ou tarefas que envolvam: identificar algum conceito estatístico, descrever gráficos, interpretar um resultado obtido de um procedimento estatístico, etc. Se o propósito é desenvolver o *raciocínio estatístico*, os alunos devem ser estimulados a saber explicar, por exemplo, porque é que a mediana é uma medida resistente a *outliers*, justificar um resultado, justificar a escolha de uma representação, explicar porquê ou como foram produzidos determinados resultados estatísticos, etc. Por fim, se o propósito é fomentar o *pensamento estatístico*, os alunos devem ser desafiados a criticar as conclusões de um estudo ou a generalizar os conhecimentos estatísticos obtidos, provenientes de exemplos escolares a novas situações, etc.

A literatura sobre o raciocínio humano revela que as pessoas apresentam dificuldades em raciocinar sobre assuntos abstratos, raciocinam melhor perante situações familiares, os seus conhecimentos pessoais influenciam frequentemente os seus raciocínios e tendem a basear-se apenas em algumas hipóteses na análise de uma situação (delMas, 2004). Pelo facto de a Estatística ser uma disciplina que se baseia fundamentalmente em situações contextualizadas, e muitas vezes familiares, podemos ser levados a pensar que os alunos não terão tantas dificuldades em raciocinar estatisticamente. Contudo, delMas (2004) refere que há estudos que mostram o contrário. O autor chama a atenção para cuidados a ter-se no raciocínio sobre o contexto dos problemas estatísticos, nomeadamente evitar-se o uso de raciocínios com base em experiências pessoais, conhecimentos quotidianos e métodos inadequados, que podem induzir interpretações erróneas.

DelMas (2004) também aponta algumas dificuldades no raciocínio estatístico dos alunos decorrentes da natureza abstrata do conteúdo estatístico. Para o autor, a Matemática que é usada na Estatística é uma fonte de abstração que pode originar dificuldades de raciocínio nos alunos. Por exemplo, no cálculo da média de um conjunto de dados obtém-se um valor exterior ao conjunto de dados e os alunos podem ter dificuldades em dar significado a este novo valor. O autor menciona ainda que compreender e raciocinar sobre representações gráficas, usadas para representar dados reais e na exploração de padrões nestes dados, tais como a caixa-de-bigodes (*boxplot*), o diagrama de caule-e-folhas (*steam-and-leaf plot*), etc, exige um certo nível de abstração que não é fácil para os alunos dado que as diferentes representações realçam mais umas características dos dados do que outras.

A unidade ou unidades de medida que foram utilizadas na recolha de dados, para um determinado estudo estatístico, podem constituir uma entidade abstrata ou não ser familiar para os alunos. Por vezes, há situações em que os alunos estão bastante familiarizados com o contexto do problema e consideram que, por exemplo, a "medida de interesse" tomada para o problema não se foca nos aspetos mais relevantes. Nesta situação, a familiaridade com o contexto do problema pode ser um obstáculo ao raciocínio estatístico dos alunos. Para delMas (2004), a partir do momento em que os alunos passem a distinguir um problema matemático de um problema estatístico e desenvolvam a compreensão dos conceitos e procedimentos em cada um destes contextos, os erros que possam cometer serão mais de natureza mecânica do que lógica. Garfield e Ben-Zvi (2007, 2008) indicam outros desafios no desenvolvimento do raciocínio estatístico dos alunos: o pouco à vontade com a confusão dos dados, as noções de aleatoriedade e acaso, as diferentes interpretações possíveis baseadas em premissas, e a grande necessidade de escrever, colaborar e comunicar. No ensino secundário, os professores devem estar preparados para desenvolver o raciocínio estatístico dos alunos, ajudando-os a saber como abordar tarefas que envolvam aprofundamento dos conceitos e a realizar análises de dados com interpretação e comunicação dos resultados (Friel, Connor & Mamer, 2006; Shaughnessy, 2007).

Apresenta-se, em seguida, uma outra conceptualização das noções de literacia, raciocínio e pensamento estatístico a partir de Ben-Zvi e Garfield (2004) e de Garfield e Ben-Zvi (2007), que é também a conceptualização adotada no presente estudo. Para estes autores: a literacia estatística “inclui capacidades básicas e importantes que podem ser usadas na compreensão da informação estatística ou resultados de investigações” (Ben-Zvi & Garfield, 2004, p. 7). Por exemplo, os alunos serem capazes de organizar dados e usar e interpretar diferentes representações quando trabalham com estes dados; demonstrarem compreensão da terminologia estatística e dos conceitos estatísticos básicos; compreenderem a probabilidade como medida de incerteza, podendo também abarcar crenças e atitudes dos alunos (Garfield & Ben-Zvi, 2007).

Por sua vez, o raciocínio estatístico “pode ser definido como o modo como as pessoas raciocinam com ideias estatísticas e dão sentido à informação estatística” (Ben-Zvi & Garfield, 2004, p. 7). O raciocínio estatístico engloba, por exemplo, os alunos serem capazes de compreender e interpretar estatísticas sumárias de dados, estabelecer ligações entre conceitos estatísticos (por exemplo, entre medidas de localização e medidas de dispersão) e entre diferentes representações estatísticas ou até relacionar ideias estatísticas com o acaso. Abrange também lidar com o processo investigativo e compreender noções estatísticas chave, nomeadamente (i) *variabilidade* (Burrill, 2008; Garfield & Ben-Zvi, 2007) – identificar, controlar, medir, analisar e explicar a variabilidade em situações diversificadas, em particular, através da capacidade de *transnumerar* com representações gráficas para chegar a um melhor entendimento dos dados; (ii) *dados e amostragem* – tipos de dados, opções de recolha de dados e o papel do contexto dos dados na amostragem (Burrill & Biehler, 2011); (iii) *associação e modelação* – analisar dados bivariados e usar a regressão, não deixando de realizar interpretações à luz do contexto dos dados (Burrill & Biehler, 2011); e (iv) *inferência* – entender a relação entre amostras e populações e ter noção da influência da qualidade dos dados nos resultados (Burrill & Biehler, 2011).

Num patamar mais elevado, “os pensadores estatísticos são capazes de criticar e avaliar resultados de um problema resolvido ou de um estudo estatístico” (Ben-Zvi & Garfield, 2004, p. 7). O pensamento estatístico envolve adicionalmente “uma compreensão do porque e como é que as investigações são conduzidas e as grandes ideias subjacentes às investigações estatísticas” (Ben-Zvi & Garfield, 2004, p. 7). Deste modo, o pensamento estatístico implica, por exemplo, os alunos demonstrarem compreender a natureza da variação estatística e da amostragem; saber em que circunstâncias e como utilizar métodos de análise dos dados e modelos estatísticos; compreender e utilizar o contexto da situação problemática em estudo no planeamento e na avaliação de investigações, na interpretação dos resultados e na comunicação das conclusões; reconhecer e compreender a importância de todas as fases de um processo investigativo, que incluem a formulação de questões, a recolha de dados, a escolha de métodos de análise de dados e realização da análise e a interpretação e apresentação dos resultados; demonstrar conhecimento das teorias que estão na base dos processos e métodos estatísticos e compreender as limitações que se encontram na Estatística e na inferência estatística (Ben-Zvi & Garfield, 2004; Garfield & Ben-Zvi, 2007). O pensamento estatístico corresponde ao modo como os profissionais estatísticos pensam (Ben-Zvi & Garfield, 2004; Garfield & Ben-Zvi, 2007; Wild & Pfannkuch, 1999).

Garfield e Ben-Zvi (2007) também assumem que há aspetos comuns entre literacia, raciocínio e pensamento estatístico, mas, ainda assim, acham que se pode considerar que existe uma hierarquia entre estas três capacidades. A literacia estatística fornece as bases para o raciocínio e o pensamento estatístico. Num nível intermédio está o raciocínio estatístico e, por fim, no patamar mais elevado desta hierarquia encontra-se o pensamento estatístico que inclui pensamento de ordem superior ao do raciocínio estatístico e consequentemente ao da literacia estatística, tal como mostra a figura 3 (que também inclui uma súmula de cada noção). Em sintonia com Garfield e Ben-Zvi (2007), Gattuso e Ottaviani (2011) acrescentam que essas três capacidades podem ser distinguidas através do nível de compreensão que as pessoas revelam relativamente a conceitos e ferramentas estatísticas e da capacidade de estabelecer conexões entre eles.

Figura 3: Relações e hierarquia entre literacia, raciocínio e pensamento estatísticos (adaptado de Garfield & Ben-Zvi, 2007, p. 381).

2.2 Ensino e aprendizagem de conceitos estatísticos

Esta secção inclui uma revisão de literatura teórica e empírica relativamente aos seguintes conceitos estatísticos centrais do ensino secundário, de acordo com o programa de Matemática A[7] (ME, 2001) e com algumas orientações internacionais (NCTM, 2007; Franklin et al., 2007): *Variabilidade*, *População e amostra*, *Medidas de localização e de dispersão* e *Dados bivariados e regressão linear*. Esta revisão tem também em consideração a investigação atual em Educação Estatística (e.g., Burrill, 2008; Burrill & Biehler, 2011; Garfield & Ben-Zvi, 2008; Shaughnessy, 2007).

[7] Programa de Matemática A do Ensino Secundário (ME, 2001) – programa em vigor no momento da realização do estudo (ver secção 2.3 e anexo 5).

2.2.1 Variabilidade

A Estatística é uma disciplina que serve para dar a outros campos de estudo um conjunto coerente de ideias e ferramentas para lidar com dados e com a variabilidade presente nos dados (Cobb & Moore, 1997; Shaughnessy, 2007). Na literatura sobre Educação Estatística, a variabilidade é considerada a componente fundamental do pensamento estatístico, sendo indiscutível a sua presença nos currículos escolares (Franklin et al., 2007; Reading & Shaughnessy, 2004; Scheafter, 2006; Shaughnessy, 2007; Wild & Pfannkuck, 1999). É também destacado o papel da variabilidade na conexão entre Estatística e Probabilidades, nomeadamente através da exploração de tarefas que incluam dados e acaso e que recaiam, em particular, sobre raciocínios que envolvam proporções (Reaburn, 2012; Torok & Watson, 2000).

De acordo com Reading e Shaughnessy (2004), o termo variabilidade "significa a caraterística de uma entidade que é observável" e o termo variação é tido como "descrição ou medição dessa caraterística" (p. 201). Ainda segundo os mesmos autores, raciocinar sobre variação diz respeito a lidar com os processos cognitivos envolvidos na descrição do fenómeno observado quanto à variabilidade. Há autores que, nas suas publicações, fazem referência à distinção entre variabilidade e variação, introduzida por Reading e Shaughnessy (2004); porém, nos trabalhos que desenvolveram optaram por usar a variabilidade como termo que abrange as duas ideias, tal como é o caso de Garfield e Ben-Zvi (2008), Makar e Confrey (2005), Sánchez, Silva e Coutinho (2011), entre outros. Esta situação deve-se, essencialmente, ao facto de considerarem que a distinção referida não teve um eco significativo na investigação em Educação Estatística.

Considerar a variação comporta não só conhecer uma definição ou fórmula (*what is it*), como usá-la de modo procedimental ou como ferramenta (*how to use it*), mas também entender e explicar o seu propósito (*why it is useful*) dentro de um contexto (Makar & Confrey, 2005). Estudar a variabilidade implica: (i) entender a sua conexão com os conceitos e propriedades dos conceitos relacionados com ela (variável, incerteza, distribuição, média, amplitude, amplitude interquartil, desvio-padrão); (ii) aprender a descrevê-la, ponderar a possibilidade de a reduzir, de representá-la e fazer previsões sobre um conjunto de dados; (iii) saber analisá-la em representações gráficas ou na comparação de gráficos ou ainda na comparação de dois ou mais conjuntos de dados (Garfield & Ben-Zvi, 2008; Shaughnessy, 2006). A análise gráfica deve ir além de se realçar um ou outro dado individual, ou de se identificar conjuntos de dados iguais, ou com uma certa especificidade no conjunto, para se chegar à visualização de caraterísticas globais do conjunto de dados, com o apoio das medidas e análise da forma da distribuição (Friel et al., 2006). Adicionalmente, analisar a variabilidade em representações bivariadas comporta identificar eventuais padrões e modelos, e realizar previsões de resultados (Engel & Sedlmeier, 2011; Garfield & Ben-Zvi, 2005, 2008).

Shaughnessy (2007) considera que os três níveis de compreensão de gráficos identificados nos trabalhos de Curcio e colegas (por exemplo, em Friel, Curcio & Bright, 2001), há muito reconhecidos na literatura, são capacidades básicas importantes que devem ser desenvolvidas nos alunos. Esses três níveis de compreensão são: (i) *Ler os dados* (identificar componentes e falar a linguagem dos dados); (ii) *Ler entre os dados* (compreender relações entre tabelas, gráficos e dados; dar sentido à informação presente no gráfico); e (iii) *Ler para além dos dados* (interpretar a informação e responder a questões sobre o gráfico; realizar inferências a partir das representações; reconhecer quais os gráficos mais adequados para determinados dados e contextos). Shaughnessy (2007) acrescenta a essa lista, um quarto nível, *Ler por detrás dos dados*, para que, na compreensão de gráficos estatísticos, se atenda especificamente à variabilidade, ou seja, procurando-se as suas causas e relações nos dados e também a conexão entre contexto e dados. Estes quatro níveis – *Ler os dados*, *Ler entre os dados*, *Ler para além dos dados* e *Ler por detrás dos dados* – são relevantes para o desenvolvimento do raciocínio estatístico dos alunos (Shaughnessy, 2007).

A compreensão da variabilidade estatística e de medidas de variabilidade (amplitude, amplitude interquartil, desvio-padrão, variância) é necessária para o entendimento do conceito de distribuição, de distribuições de amostragem e da inferência estatística (Burrill & Biehler, 2011; delMas & Liu, 2005; Garfield & Ben-Zvi, 2008). Vários estudos revelam que os alunos de diversos níveis escolares têm dificuldades em interpretar gráficos, em reconhecer quais os mais adequados para um dado conjunto e seu contexto e também em analisar a variabilidade dos dados através deles, capacidades essenciais no desenvolvimento da literacia, raciocínio e pensamento estatístico (Friel et al., 2001; Shaughnessy, 2007). Groth (2013), por exemplo, detetou dificuldades nalguns dos seus alunos de licenciatura (futuros professores do ensino básico) na análise de gráficos não convencionais produzidos por alunos do ensino básico. A este respeito, o autor recomenda que na formação inicial os futuros professores melhorem a capacidade de analisar gráficos, atendendo às suas qualidades e eventuais aspetos inadequados. Realçou especificamente a necessidade destes professores saberem como relacionar gráficos e conhecerem os que devem ser utilizados no ensino para apoiar a compreensão do diagrama de extremos e quartis.

No estudo conduzido por Oliveira e Henriques (2014), com futuros professores dos ensinos básico e secundário, a componente variação ou variabilidade do pensamento estatístico foi a única componente que não emergiu na análise dos relatórios reflexivos de futuros professores sobre episódios de aula de uma investigação estatística. Num outro estudo com quatro professores do ensino secundário que tinha também como objetivo perceber como é que os professores ensinavam a variabilidade a partir de análises de gráficos, González (2013) constatou que estes participantes, de uma maneira geral, apresentaram poucas situações nas suas respostas escritas, parecendo não conhecer estratégias diversificadas.

2.2.2 População e amostra

O reconhecimento da necessidade de dados e de que os dados são números com um *contexto* é uma componente do pensamento estatístico a ter em conta no contexto escolar (Burrill & Bihler, 2011; Scheaffer, 2006; Wild & Pfannkuck, 1999). A metodologia adequada para obtenção de dados deveria garantir representatividade e a variabilidade presente na amostra deveria refletir a variabilidade da população. Além desta variabilidade desejada há a considerar a variabilidade indesejada que necessita de ser controlada (no processo de seleção de unidades amostrais, na sua medição ou observação) e que a Estatística tenta colocar ao seu serviço para obter dados, observando-os ou produzindo-os. Eis alguns exemplos de fontes de variabilidade a controlar (Franklin et al., 2007), sem se colocar em causa a representatividade da amostra (preferencialmente, com custos mínimos): *variabilidade de medição*, em que medições repetidas no mesmo indivíduo variam; *variabilidade natural*, que se refere a uma variabilidade inerente na natureza. Por exemplo, as pessoas naturalmente têm diferentes alturas, aptidões e capacidades, diferentes opiniões e respostas emocionais e, consequentemente, quando medimos qualquer um destes aspetos sabemos que vamos ter uma grande variabilidade; *variabilidade induzida* (por fatores), por exemplo, há estudos que comprovam que certas qualidades de milho crescem melhor num clima do que noutro; que um fertilizante é mais eficaz do que outro, etc. Um estudo experimental cuidadosamente concebido pode ajudar a determinar os efeitos de diferentes fatores e eventualmente estabelecer relações causais; *variabilidade da amostragem*, por exemplo, numa sondagem política a proporção de votantes de uma amostra para se estimar a proporção desconhecida de todos os votantes que apoiam determinado candidato varia de amostra para amostra e designa-se por variabilidade da amostragem. No processo de amostragem, essa variabilidade tem também tendência para diminuir à medida que o tamanho da amostra aumenta.

No ensino, há que desenvolver o conhecimento estatístico dos alunos sobre populações, amostras e amostragem em torno de noções adequadas, partindo de ideias mais simples para as mais complexas, para que estes tenham gradualmente mais presente características de boas amostras e reconheçam mais facilmente amostras enviesadas (Garfield & Ben-Zvi, 2008; Rubin et al., 1990; Sharma, 2005). Por exemplo, o valor de uma medida estatística proveniente de uma amostra não tem de corresponder necessariamente ao valor do respetivo parâmetro populacional; contudo, este valor referente a uma amostra pode fornecer uma boa informação sobre o respetivo parâmetro (Garfield & Ben-Zvi, 2008; Rubin et al., 1990). Outro aspeto a ter em conta é que a variação que se encontra de amostra para amostra é uma consequência do processo de amostragem e não devido ao erro. Além disso, o tamanho de uma amostra representativa não se define obrigatoriamente através de uma percentagem particular da população; existem boas amostras representativas que correspondem efetivamente a uma pequena percentagem da sua população (Garfield & Ben-Zvi, 2008).

Rubin et al. (1990) realizaram um estudo com doze alunos do ensino secundário para procurarem perceber que conceções estes detinham acerca da representatividade e variabilidade de amostras. Os alunos revelaram dificuldades em descrever de uma forma adequada as relações entre populações e amostras que lhes foram apresentadas. Os autores consideram que essas dificuldades parecem estar relacionadas com a falta de experiência dos alunos em raciocinar sobre amostras e em fazer inferências e devido à tendência de lidar com a Estatística do mesmo modo como trabalham a Matemática, nomeadamente, em busca de respostas certas ou exatas.

Garfield e Ben-Zvi (2008), Rubin et al. (1990) e Sharma (2005) também alertam para a importância de se proporcionar experiências diversificadas na aula sobre recolha de dados e sobre o processo de amostragem. Através de tais experiências, os professores são capazes de ajudar os alunos a realizar uma aprendizagem em que aprofundem os conceitos e a fornecer descrições e explicações adequadas em torno da variabilidade da amostra, mesmo quando recorrem a argumentos que retiram das suas experiências pessoais.

2.2.3 Medidas de localização e de dispersão

Na literatura em Educação Estatística, média, moda e mediana são denominadas como medidas de tendência central ou como medidas de localização, ou ainda como medidas de centro. Estas medidas no contexto escolar devem ser desenvolvidas ao nível conceptual e processual (Ben-Zvi & Garfield, 2004; Friel, 1998), em particular no ensino secundário (Langrall & Mooney, 2010). Vários autores recomendam que se proporcionem tarefas diversificadas através das quais as medidas de localização possam ser comparadas entre si, em particular, relativamente às suas propriedades e também relacionadas com outras noções e representações gráficas (e.g., Friel, 1998; Friel et al., 2006; Shaughnessy, 2007; Utts, 2003). Consideram também que é essencial que os alunos ganhem experiência para saber quando e como usar as medidas de localização na análise de dados e para saber explicar o que é que elas transmitem sobre os dados. Assim sendo, há autores que sugerem que, no ensino da Estatística, as medidas de localização surjam associadas à variabilidade, devendo ser analisadas em conjunto na caraterização de conjuntos de dados e também na comparação desses conjuntos (Garfield & Ben-Zvi, 2008; Konold & Pollatsek, 2004; Langrall & Mooney, 2010; Shaughnessy, 2007).

Vários autores da área da Educação Estatística (e.g., Garfield & Ben-Zvi, 2008; Konold & Pollatsek, 2004; Shaughnessy, 2007) recomendam que o desenvolvimento das medidas de centro, no ensino, se realize de uma forma gradual, a partir de ideias intuitivas iniciais dos alunos tais como *meios* ou *a maior parte*, para conceitos mais elaborados como *partilha equitativa*, *valor típico*, *ponto de equilíbrio*, relevantes para a introdução das medidas de centro. Recomendam ainda que passe também por outros mais sofisticados como *redutor de dados* ou *sinal no meio do ruído*. Em particular, consideram que a média como *ponto de equilíbrio* e *partilha equitativa* são conceitos facilmente associáveis ao algoritmo da média. A abordagem de nivelamento (*partilha equitativa*) pode conduzir à visualização da média como *ponto de equilíbrio*. E a média como *ponto de equilíbrio* deve permitir entender, por exemplo, por que é nula a soma dos desvios à média.

Do ponto de vista conceptual, a média poderá ser aprofundada quando visualizada como ponto de equilíbrio (*balance point*) de uma distribuição ou associada à partilha equitativa (*fair-share*) e também quando analisada quanto à sua resistência face à presença de eventuais *outliers* ou valores extremos, mesmo em situações de médias ponderadas. Ao considerar-se uma distribuição de dados quantitativos, no primeiro modelo (*balance point*), a média opera como um fulcro de uma balança e é o valor sobre o qual os dados ficam equilibrados (ver figura 4), dado que o valor absoluto da soma das diferenças entre cada dado à esquerda da média e a média e o valor absoluto da soma das diferenças entre cada dado à direita da média e a média são iguais. No segundo modelo (*fair-share*), o valor da média coincide com o valor que diz respeito à redistribuição equitativa dos valores dos dados sobre cada valor que a variável estatística assuma. Estas duas abordagens destacam a média como um valor representativo de um conjunto de dados (Friel, 1998; Konold & Pollatsek, 2004).

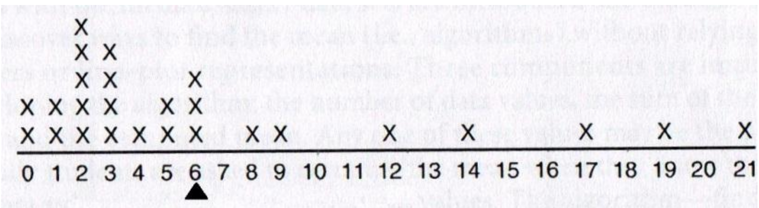


Figura 4: Gráfico de pontos (com frequências unitárias *x*; Friel, 1998, p. 212).

Associar média e mediana a *sinal no meio do ruído* (sendo estas associadas, respetivamente, ao desvio-padrão e à amplitude interquartil) é uma ideia importante e útil na comparação de dois conjuntos de dados. Ver a média e a mediana como *redutores de dados* pode não ser acessível para muitos alunos, uma vez que, com o processo de redução (através de um cálculo ou estimativa sobre uma representação tabular ou gráfica), pode acontecer que alguma informação detetada previamente tenha deixado de estar presente (Garfield & Ben-Zvi, 2008; Shaughnessy, 2007).

Contudo, no contexto escolar, as experiências proporcionadas aos alunos com as noções de centro (média, moda e mediana) ficam muitas vezes reduzidas à esfera computacional. Uma aprendizagem centrada no cálculo da média não permite desenvolver o conceito e pode mais facilmente inibir a capacidade de relacionar a média com outros conceitos (Pollatsek, Lima & Weel, 1981). Vários estudos que analisaram a aprendizagem das medidas de localização sugerem que alunos de diferentes níveis escolares, alunos universitários e professores enfrentam dificuldades com essas medidas (Ben-Zvi & Garfield, 2004; Groth & Bergner, 2006; Shaughnessy, 2007). Em particular, alguns estudos nos anos 80 e 90 realizados com alunos do ensino básico e secundário mostraram as suas dificuldades em resolver questões que envolviam propriedades básicas da média e o uso da média na comparação de conjuntos de dados de tamanhos diferentes (Garfield & Ben-Zvi, 2008).

Hardiman, Well e Pollatsek (1986) realizaram um estudo relacionado com o conceito de média, mais especificamente, sobre o conceito de equilíbrio e de como os alunos relacionam informação sobre quantidade de pesos e a sua distância ao fulcro de uma balança, para conjeturar se esta ficaria equilibrada ou não e, em caso negativo, identificar o lado para o qual se inclinaria. Participaram no estudo alunos universitários que realizaram inicialmente um teste que incluía várias situações problemáticas sobre equilíbrio e que aumentavam progressivamente de nível de dificuldade. Alguns destes alunos foram selecionados para participar em sessões individuais em que tiveram oportunidade de, simultaneamente, analisar diversas situações sobre equilíbrio e de ver o resultado de cada previsão que realizaram antecipadamente numa balança. Neste contexto, os alunos teriam de explicar os seus raciocínios ao longo das diferentes tarefas e o investigador/entrevistador procurava perceber quando é que estes alunos chegavam a alguma regra geral, nomeadamente, se percebiam a importância da contabilização das distâncias para obter uma previsão correta. Os autores deste trabalho concluíram que todos os alunos que participaram nas sessões desenvolveram a sua compreensão sobre equilíbrio. Com este estudo, estes autores também aproveitaram para reforçar a ideia do valor deste conhecimento no contexto educacional, ou seja, em promover o entendimento concetual da média.

Num estudo com alunos dos ensinos básico e secundário, foram usados quatro pares de gráficos (figura 5) em entrevistas (Watson, 2001; Watson & Shaughnessy, 2004). Os dados e contexto associados a cada gráfico (que dizem respeito a turmas de alunos) não têm qualquer relação com os participantes do estudo. Cada par de gráficos representava uma turma diferente e, em cada gráfico, cada quadrícula indicava o número de respostas corretas dadas por cada aluno da turma num teste. Nas entrevistas, os alunos participantes teriam de identificar e explicar qual a turma com melhores resultados em cada par. Os autores queriam perceber quais as estratégias utilizadas pelos alunos para responder à questão, em particular, de que forma usavam a média. Para decidir se as duas turmas teriam resultados iguais ou se alguma delas teria melhores resultados, poucos alunos utilizaram a média, nomeadamente, no par que representava duas turmas de dimensões diferentes. A maioria desses alunos lembrava-se do algoritmo; ainda assim, alguns fizeram erros quando usaram a calculadora para realizar o cálculo. Destes alunos que referiram a média, poucos a associaram à ideia de nivelamento e a maioria teve dificuldade em explicar por que era apropriado usar a média nas comparações. Identificou-se, ainda, que três alunos não usaram a média mas, através da observação dos gráficos e de referências a certas proporções, revelaram entender o cerne da questão. Watson e Shaughnessy (2004) consideram que a exploração e discussão em sala de aula de tarefas que envolvam comparação de conjuntos de dados e atendimento à variabilidade através do confronto entre aspetos visuais e numéricos constituem oportunidades valiosas para os alunos desenvolverem o seu raciocínio estatístico. Os autores salientam a importância dos professores estarem preparados para promover essas explorações e discussões nas aulas com vista ao alargamento do conhecimento dos alunos.

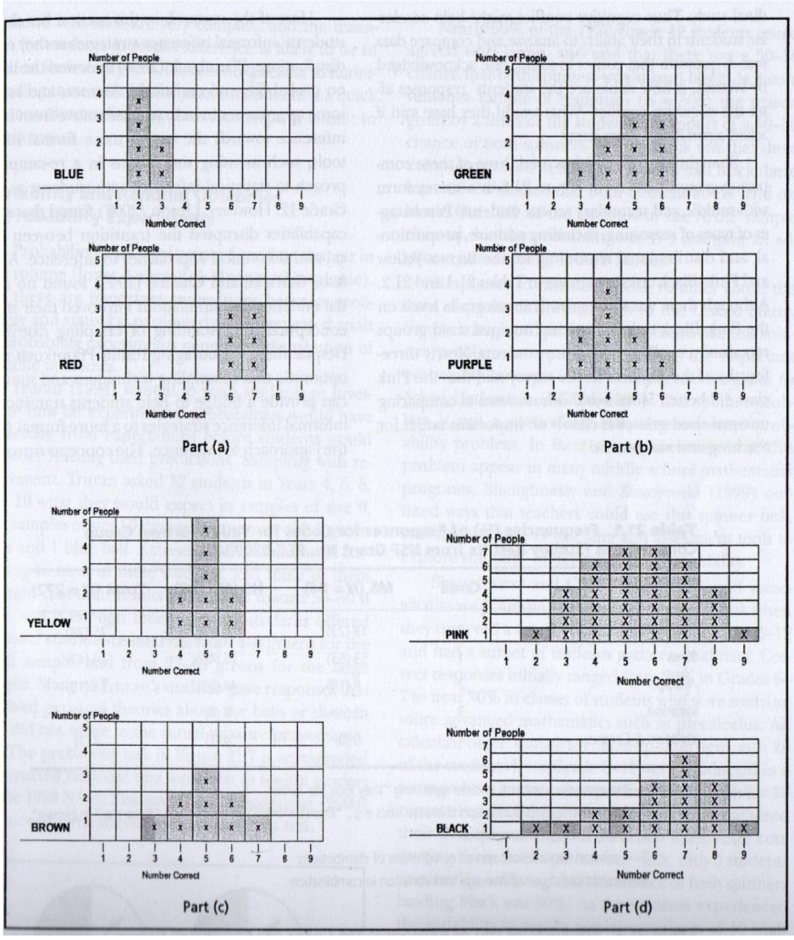


Figura 5: Quatro pares de gráficos para comparar conjuntos de dados (Watson, 2001).

Boaventura e Fernandes (2004) realizaram um questionário que contou com a participação de 181 alunos do ensino secundário (de três turmas do 12.º ano de escolas diferentes de uma mesma cidade) e que incluía algumas questões diversificadas que implicavam a utilização das medidas de tendência central (média, mediana e moda) nas suas resoluções. As situações apresentadas estavam ao nível do 3.º ciclo e do 10.º ano. Os autores concluíram que a maioria dos alunos revelou dificuldades em lidar com a noção de média ponderada; em determinar um conjunto de dados sendo conhecidos os valores de média, mediana, moda e o tamanho da amostra; em explicar o significado do resultado de uma média, moda ou mediana de uma situação contextualizada; e ainda, em identificar média, moda e mediana em três distribuições representadas graficamente. Os autores são da opinião que o conhecimento das dificuldades e dos raciocínios dos alunos em torno das medidas de tendência central é um assunto relevante para reflexão e que também deverá contribuir para que os professores aperfeiçoem a sua atuação no sentido de ajudar os alunos a melhorar o seu conhecimento de conceitos elementares de Estatística e os seus desempenhos.

De seguida, apresentam-se resultados de algumas investigações relativamente ao conhecimento revelado por professores e futuros professores com respeito a medidas de tendência central. Num estudo cujos participantes eram futuros professores e professores dos ensinos básico e secundário, cerca de 42% errou numa questão que envolvia uma média ponderada (Callingham, 1997). Num outro estudo com 46 futuros professores do ensino básico, que tinham frequentado duas disciplinas de Matemática elementar e uma disciplina introdutória de Estatística na universidade, a maioria revelou possuir um conhecimento das medidas de tendência central focado essencialmente nalgumas regras, algoritmos e procedimentos, necessitando de o aprofundar a nível concetual (Groth & Bergner, 2006). De acordo com os dados deste estudo, os autores identificaram e descreveram quatro níveis de conhecimento de conteúdo do professor baseados na taxonomia SOLO (*The Structure of the Observed Learning Outcome*, de Biggs e Collis, 1982), os quais se apresentam, em seguida, do mais elementar para o mais complexo, indicando também o número de participantes classificados em cada um deles:

> Nível *uni-estrutural* – a única estratégia usada nas respostas foi a indicação da definição de cada medida, quando era pedido a comparação e contrastação das três medidas de localização (n=8).
>
> Nível *multi-estrutural* – as respostam incluíam a definição de cada medida e indicavam, de uma forma superficial, que as três medidas eram ferramentas que podiam ser usadas para analisar um conjunto de dados (n=21).
>
> Nível *relacional* – as respostas incluíam o reconhecimento de que moda, média e mediana descrevem o centro ou estão próximas do meio do conjunto de dados ou o que era típico nos dados (n=13).
>
> Nível *abstrato estendido* – as respostas iam além do nível anterior, incluíam discussões de quando uma medida poderia ser mais útil do que outra (n=3).

Todas as respostas dos participantes foram categorizadas num desses níveis, exceto uma, por não se enquadrar claramente apenas num dos níveis. Os autores concluíram que na formação inicial os futuros professores devem trabalhar as noções de média, moda e mediana e suas relações com recurso a tarefas diversificadas, para que depois sejam capazes de selecionar ou até produzir tarefas para a aula que permitam desenvolver as aprendizagens desejadas nos seus alunos.

Num estudo mais recente de Watson e Callingham (2013), realizado na Austrália, foi analisado o *conhecimento pedagógico de conteúdo*[8] relativamente ao conceito de média, a partir de um questionário a 26 professores do ensino básico e secundário. Foram-lhes colocadas três questões abertas: a primeira implicando a tomada de decisões no ensino da média, e a segunda e a terceira envolvendo a avaliação de duas respostas inadequadas de alunos a uma situação específica que envolvia recalcular o valor de uma média devido à necessidade de se eliminar alguns dados (sendo conhecidos o tamanho da amostra e o valor da média). As autoras concluíram que a maioria dos professores participantes conseguiu listar uma sequência de tópicos relacionados com a média embora estabelecendo poucas relações entre eles. Na avaliação das respostas inadequadas de dois alunos, quase todos os participantes não tiveram dificuldade em apontar os erros no cálculo da média numa dessas respostas, fornecendo explicações que tinham a ver com a composição da fórmula da média. Nesta questão cerca de 50% dos professores indicou algumas estratégias que poderiam ajudar os alunos a alcançar uma resposta adequada. A maioria dos professores também comentou a resposta inadequada de outro aluno, apresentando explicações que tinham como referência a fórmula da média ou apontando algum elemento dessa fórmula. Apenas seis professores apresentaram sugestões que apontavam alguns exemplos e que poderiam ser eventualmente usados para o entendimento da questão. Estes dois últimos estudos referidos enfatizam a importância dos professores de cada nível de ensino aprenderem a articular os conhecimentos estatísticos que possuem (nomeadamente, os relacionados com medidas de centro) e serem capazes de mobilizá-los adequadamente no ensino (Groth & Bergner, 2006; Watson & Callingham, 2013).

[8] Esta noção é desenvolvida no capítulo 3 (O conhecimento didático do professor em Estatística).

O desvio-padrão é uma medida de variabilidade cuja ênfase no ensino está, muitas vezes, no uso de uma fórmula e na repetição do procedimento (delMas & Liu, 2005). Das medidas de dispersão ou variabilidade – amplitude total, amplitude interquartil e desvio-padrão – a amplitude total é a mais acessível para os alunos e o desvio-padrão a mais complicada das três. A maior dificuldade detetada nos alunos não reside nos cálculos mas na interpretação dos resultados numéricos destas medidas e na sua compreensão em representações. Há a necessidade de se promover o conhecimento concetual das medidas de variabilidade relacionando-as simultaneamente com outros conceitos (Garfield & Ben-Zvi, 2005, 2008).

Para se promover uma compreensão significativa do desvio-padrão será necessário estudar um conjunto de conceitos estatísticos, nomeadamente aqueles a partir dos quais este se constrói: *distribuição* (que, por sua vez, requer um entendimento do conceito de variável e coordenação dos seus valores e densidades de uma distribuição quando representada graficamente); *média* (deve desenvolver-se no aluno outras noções além do algoritmo, por exemplo, média como *ponto de equilíbrio*) e o conceito de *desvio à média*. Desta forma, os alunos poderão desenvolver a capacidade de refletir sobre propriedades do desvio-padrão, por exemplo, a sua resistência face à presença de algum valor extremo na distribuição e, em representações gráficas, a análise de como é que os valores de uma variável e suas respetivas frequências poderão, independente ou conjuntamente, afetar o seu valor (delMas & Liu, 2005; Garfield & Ben-Zvi, 2008). Analisar se há muita ou pouca variabilidade nos dados, a amplitude e amplitude interquartil em *boxplots* e o que pode fazer com que essas medidas variem, são também atividades importantes para o desenvolvimento desses conceitos nos alunos (Garfield & Ben-Zvi, 2008).

2.2.4 Dados bivariados e regressão linear

O estudo de dados bivariados envolve análises de variáveis individuais e sua variabilidade, e da forma das distribuições bivariadas em gráficos (Garfield & Ben-Zvi, 2008). O conhecimento de que a existência de correlação não implica que haja uma relação causal entre os dados bivariados em estudo (Friel et al., 2001) não deve ser descurado no ensino, bem como uma inferência deste tipo só poderá ser realizada através de uma experiência estatística (Estepa & Batanero, 1996; Garfield & Ben-Zvi, 2008; Pestana & Velosa, 2010).

Garfield e Ben-Zvi (2008) sugerem que atividades de estabelecimento de conexões entre valores de coeficientes de correlação e diagramas de dispersão podem permitir que os alunos desenvolvam um melhor entendimento dos diferentes níveis de covariação e dos fatores que influenciam o maior ou menor valor do coeficiente de correlação. Por exemplo, um resultado de correlação linear alto não implica, necessariamente, a validade do modelo de regressão linear, devendo examinar-se cuidadosamente representações gráficas dos dados, como o diagrama de dispersão, uma vez que o coeficiente de correlação linear é uma medida muito influenciada por *outliers*.

Em sintonia com esta ideia, Anscombe, num artigo publicado em 1973, discute a importância dos gráficos em análises de regressão. Nesse trabalho, apresenta quatro conjuntos de dados, com coeficientes de correlação linear próximos de 0.82, cada um representado num diagrama de dispersão (ver figuras 6 e 7) com a sua respetiva reta de regressão (y=0.5x+3), que é a mesma nas quatro situações. De acordo com o autor, a análise gráfica de cada um dos gráficos permite, de certo modo, avaliar a associação entre as variáveis em causa. Nomeadamente, o gráfico A parece descrever uma relação linear, o B curvilínea, possivelmente quadrática, e os gráficos C e D ilustram o efeito de um *outlier*. Anscombe (1973) alerta que estas análises deverão ser complementadas com o contexto. Por conseguinte, um coeficiente de correlação por si só não deve ser usado para se retirarem conclusões sobre uma eventual associação linear entre duas variáveis quantitativas.

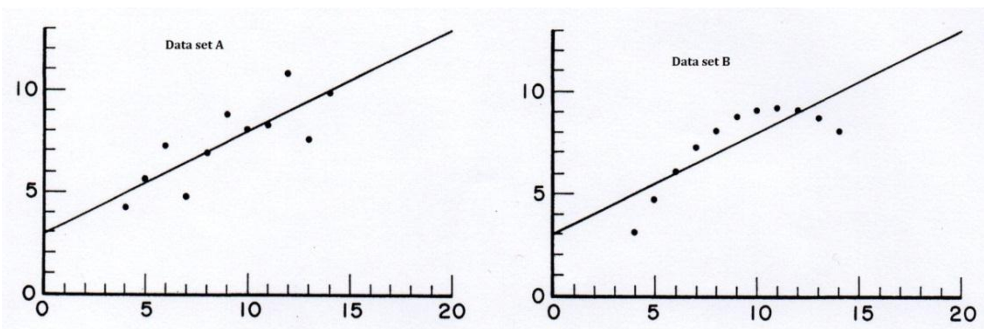


Figura 6: Gráfico A e Gráfico B (Anscombe, 1973, p. 19).

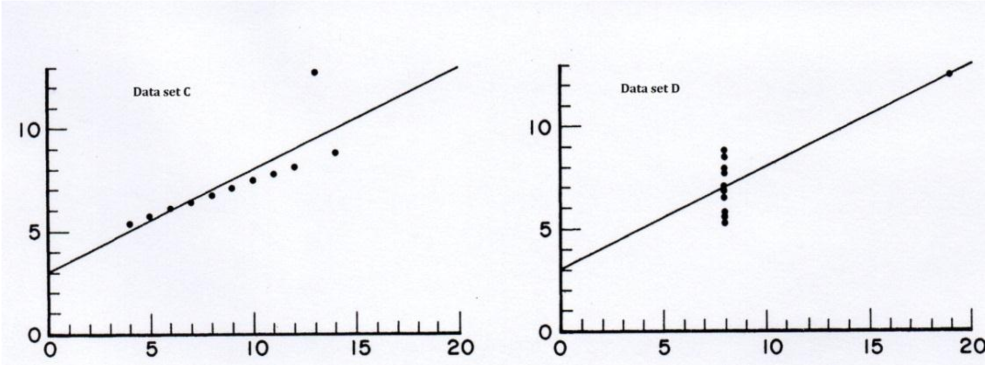


Figura 7: Gráfico C e Gráfico D (Anscombe, 1973, p. 20).

Vários autores (e.g., Shaughnessy & Chance, 2005) também referem exemplos de conjuntos de dados para os quais a afirmação de que a correlação positiva entre duas variáveis traduzida como *assim que uma delas aumenta, a outra também aumenta*, nem sempre é verdadeira. Para esse caso, uma afirmação mais precisa é: *valores acima da média de uma das variáveis correspondem a valores acima da média da outra variável*. Tal pode ser observado no gráfico seguinte (figura 8), respeitante à possível associação entre o preço (y) e a área (x) de uma amostra de 25 casas, com os dados que se enquadram no quadrante II.

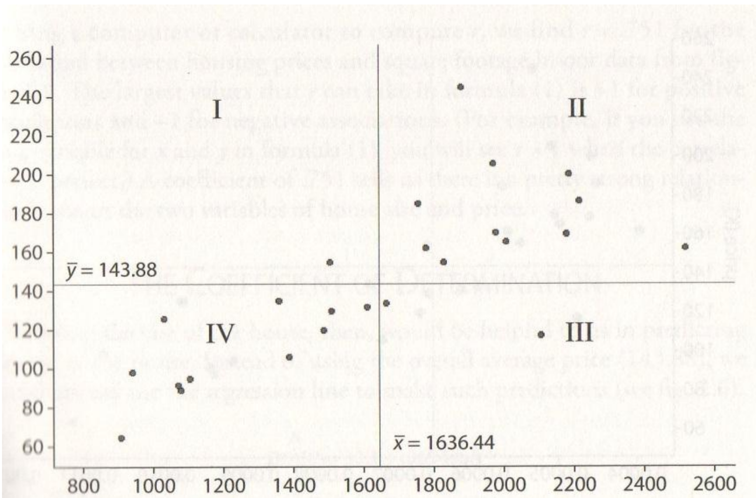


Figura 8: Diagrama de dispersão (Shaughnessy & Chance, 2005, p. 11).

Esse conhecimento mais pormenorizado e associado ao centro de gravidade de uma distribuição poderá contribuir para uma melhor apreciação da variação local e global dos dados. Adicionalmente, poderá ainda ajudar no entendimento da fórmula e resultado do coeficiente de correlação e na compreensão deste conceito como medida para a força de associação de duas variáveis (Franklin et al., 2007; Shaughnessy & Chance, 2005). A tecnologia pode proporcionar a visualização de representações diversificadas, com eventual exploração e manipulação dos dados, e o estabelecimento de conexões entre elas. Esta poderá também, por um lado, promover um entendimento mais profundo dos conceitos envolvidos e, por outro, ser um apoio significativo à análise de dados (Franklin et al., 2007; Garfield & Ben-Zvi, 2008).

Atividades de natureza investigativa que envolvam, em particular, dados bivariados, podem permitir que os alunos desenvolvam um conjunto de capacidades que refletem a constante interação entre dados e o seu contexto, nomeadamente: *analisar o problema*, procurando padrões e relações nos dados; *implementar e avaliar o método escolhido*, usando representações convencionais e não convencionais, cálculos e modelos; *procurar e estabelecer relações*, com as representações numéricas e gráficas, identificando aspetos comuns e distintos das diversas componentes de análise e ligando conclusões e interpretações ao contexto; e ainda *refletir sobre os resultados*, tomando em consideração limitações da investigação, a variabilidade e o contexto dos dados (Shaughnessy, Chance & Kranendonk, 2009, p. 3). Verifica-se, porém, que a investigação com respeito à compreensão que o professor revela sobre correlação e regressão é escassa (Engel & Sedlmeier, 2011; Shaughnessy, 2007). Mesmo quanto à compreensão que os alunos detêm destes tópicos, continua a haver pouca investigação.

Recentemente, o foco de trabalho de investigação tem-se centrado na forma como os alunos interpretam representações gráficas com apoio da tecnologia (através de softwares estatísticos) ou sem ela (Garfield & Ben-Zvi, 2008). Engel e Sedlmeier (2011) consideram que há, ainda, um longo caminho a percorrer, que passa pela realização de estudos empíricos nos diversos níveis de escolaridade.

Estepa e Batanero (1996), num estudo realizado com 213 estudantes que frequentavam o último ano do ensino secundário (cerca de 18 anos de idade), investigaram as suas estratégias na avaliação de correlações entre variáveis numéricas, identificando alguns erros de associação cometidos que estes cometeram. Duas das conceções identificadas foram a *local* (quando os alunos usaram apenas um parte dos dados no seus julgamentos, mas generalizaram os resultados para o conjunto inteiro) e a *causal* (quando os alunos reconheciam a associação e acreditavam que existia uma relação causal entre as variáveis).

Estepa e Cobo (1998) realizaram um estudo em que se analisou como os tópicos de correlação e regressão estavam organizados e expostos em onze manuais escolares do ensino secundário espanhol (escolhidos de um total de 28 manuais publicados no período de 1977 a 1990, pelas principais editoras espanholas). Os autores concluíram que a maioria destes manuais não apresentava os aspetos essenciais das noções envolvidas e que sobrevalorizava os cálculos.

Cobb, McClain e Gravemeijer (2003) conduziram um estudo que consistiu no estabelecimento de uma trajetória de aprendizagem com alunos do ensino básico (8.º ano) em que os alunos tiveram de analisar conjuntos de dados bivariados através de *applets* construídos especificamente para esse estudo (*Minitools software*). Os resultados sugerem que os alunos, ao analisarem especificamente a variação da variável y (*distribuição vertical*) para cada valor da variável x dos dados bivariados representados graficamente, ficaram com uma melhor perceção da variabilidade dos dados e da forma global da distribuição. Os autores também indicaram que esse entendimento poderia facilitar uma maior discussão e compreensão à volta da força e sinal da correlação e da reta que melhor se ajustava à distribuição bivariada.

Engel e Sedlmeier (2011) registam que, com frequência, os dados bivariados são trabalhados na aula como uma dependência funcional entre duas variáveis, descurando-se a variação dos dados. Sorto, White e Lesser (2011) referem que as primeiras conceções dos alunos na procura da *reta que melhor se ajusta aos dados do diagrama de dispersão* podem não ter a ver com minimizar distâncias entre pontos e a reta, por isso, apelam à necessidade de se perceber as inclinações usuais dos alunos, analisá-las do ponto de vista matemático e, posteriormente, ser trabalhada a ideia de que o melhor ajustamento ocorre com o método dos mínimos quadrados.

Na regressão linear os dados bivariados (x_i, y_i) podem ser traduzidos por duas componentes - uma componente estrutural e uma componente residual -, ou seja, “dados

= sinal[9] + ruído[10]" (p. 253). Este modelo de regressão pode ser expresso matematicamente pela fórmula $y_i = f(x_i) + e_i$, em que a primeira parcela representa a variação explicada e, a segunda, a variação não explicada. Este modelo pode permitir que se faça uma melhor apreciação da variação quando visualizada num diagrama de dispersão (Engel & Sedlmeier, 2011).

No ensino da Estatística, destaca-se a importância de se proporem tarefas que incluam dados reais, elementos sobre o seu contexto e questões que valorizem os dados (Curcio & Artz, 1996; Engel & Sedlmeier, 201; Scheaffer, 2006). Em particular, o raciocínio sobre dados bivariados deve suportar algo mais do que raciocinar sobre diagramas de dispersão, correlação, regressão e funções. Deve envolver o entendimento de ideias de estrutura e força na relação bivariada, com a análise residual e o ajuste do modelo, bem como a compreensão do papel da relação bivariada em modelos e na previsão de eventos (Garfield & Ben-Zvi, 2008). O planeamento de atividades que cubram situações diferentes e que também envolvam o recurso à tecnologia tem o potencial de desenvolver o raciocínio estatístico dos alunos nesta temática (Engel & Sedlmeier, 2011; Garfield & Ben-Zvi, 2008; Sorto et al., 2011).

2.3 Orientações curriculares para a Estatística

Nesta secção apresenta-se uma síntese de algumas orientações curriculares internacionais sobre o ensino da Estatística, nomeadamente, os *Princípios e normas para a matemática escolar* do NCTM (2007) e o relatório *GAISE* (Franklin et al., 2007), que têm uma presença vincada na revisão de literatura atual na área da Educação Estatística. Inclui-se ainda uma breve apresentação, no que diz respeito à Estatística, do *Programa de Matemática A do Ensino Secundário* (ME, 2001) em vigor no momento da realização deste estudo.

---

[9] Sinal – que pode ser igualmente designado por: variação explicada; modelo que se ajusta aos dados; estrutura.

[10] Ruído – que pode ser igualmente designado por: erro, variação não explicada; resíduos; aleatoriedade.

2.3.1 As orientações do NCTM e do relatório GAISE

As normas do NCTM (2007) referentes à Estatística, nomeadamente, sobre o tópico *Análise de dados e probabilidades* sugerem um ensino que abranja todas as fases do processo investigativo estatístico (Formulação do problema; Planificação e recolha de dados; Análise de dados; Interpretação e comunicação dos resultados) desde o pré-escolar ao 12.º ano:

> Formular questões que possam ser abordadas por meio de dados e recolher, organizar e apresentar dados relevantes que permitam responder a essas questões; Selecionar e usar métodos estatísticos adequados à análise de dados; Desenvolver e avaliar inferências e previsões baseadas em dados. (p. 52)

Para os anos do 9.º ao 12.º, as normas expressam a necessidade de aprofundar os conhecimentos acerca de como lidar com a variabilidade dos dados durante o processo investigativo, de modo a que as conclusões realizadas reflitam análises adequadas à variabilidade. Adicionalmente, os alunos devem aprender a trabalhar com o processo investigativo de uma forma cada vez mais elaborada, questionando as situações que advenham dessas fases. Por exemplo, aprendendo "a colocar questões que os ajudem a avaliar a qualidade de sondagens, de estudos de observação e de experiências controladas. (…) Deverão tornar-se consumidores entendidos, críticos e ponderados da informação e dos dados concebidos por outros" (NCTM, 2007, p. 383). Essas normas também indicam a potencialidade dos alunos aprofundarem as suas aprendizagens na Estatística ao relacionar os seus conhecimentos prévios de Matemática (gráficos e funções) com a Estatística, em particular quando trabalham com dados bivariados e quando começam a inteirar-se da ligação da Estatística com as Probabilidades, como passo importante para a aceder à inferência estatística.

Apresenta-se, de seguida, de uma forma sumária, as expetativas para a aprendizagem da Estatística mencionadas no documento do NCTM (2007) do 9.º ao 12.º ano, através do tópico *Análise de dados e probabilidades* (adaptado de NCTM, 2007, pp. 382-392):

Formular questões que possam ser abordadas por meio de dados e recolher, organizar e apresentar dados relevantes que permitam responder a essas questões

- Entender diferenças entre estudos (sondagens, estudos de observação e experiências) e suas aplicações; conhecer características de estudos bem planeados e o papel da aleatoriedade nesses estudos; - Compreender o termo variável; dados quantitativos/qualitativos e dados univariados/bivariados; - Compreender e usar representações gráficas (histogramas, diagramas de extremos e quartis e diagramas de dispersão) e numéricas elementares. Distinguir medida de um parâmetro.

Selecionar e usar métodos estatísticos adequados à análise de dados

- Usar representações gráficas e numéricas adequadas na apresentação de dados quantitativos univariados e na descrição da forma da distribuição; entender como é que transformações lineares nos dados afetam a forma, centro e dispersão. - Usar representações gráficas e numéricas no tratamento de dados quantitativos bivariados, com recurso à tecnologia. Análise residual. Discutir dados bivariados com variáveis qualitativas. Identificar tendências e proceder à modelação (encontrando funções ou transformando os dados para poderem ser modelados).

Desenvolver e avaliar inferências e previsões baseadas em dados

- Desenvolver algumas intuições em torno da amostragem. Usar simulações para explorar a variabilidade de medidas estatísticas, construir e usar distribuições amostrais. -Avaliar relatórios publicados de dados (plano, adequação da análise e validade de conclusões). Compreender algumas técnicas estatísticas no controlo da qualidade (amostragem sistemática).

Compreender e aplicar conceitos básicos de probabilidades

- Compreender as noções de espaço amostral, distribuição de probabilidades e de probabilidade condicionada. Usar simulações para criar distribuições de probabilidade. Calcular e interpretar a média. Compreender os conceitos de acontecimentos independentes e acontecimentos compostos. Calcular probabilidades.

O relatório *GAISE* (Fanklin et al., 2007) sugere também que o ensino da Estatística, além de incorporar o processo investigativo, deve também considerar a variabilidade dos dados, como controlá-la e analisá-la à medida que se vai aumentando o nível de complexidade dos conceitos e das atividades ao longo dos vários níveis de escolaridade com vista ao desenvolvimento do pensamento estatístico. Este relatório considera três níveis de desenvolvimento, um elementar, um intermédio e um outro mais avançado, que devem ser experienciados pelos alunos para que desenvolvam os seus conhecimentos estatísticos de forma adequada. Apesar de se poder estabelecer uma correspondência entre esses três níveis de desenvolvimento e níveis de escolaridade, o relatório não o especifica dessa forma; ainda assim, indica que o nível mais avançado deverá ser alcançado pelos alunos no ensino secundário. Descreve-se, de seguida, de uma forma sumária, em que consiste cada um desses níveis na *Análise de dados* e *interpretação dos resultados* (adaptado de Franklin et al., 2007, pp. 11-59):

A um *nível de desenvolvimento elementar*, os alunos devem saber distinguir dados qualitativos dos quantitativos, entender quais as medidas e representações gráficas a usar para cada tipo de variável; devem comparar dados individuais entre si, dados individuais com o seu conjunto de dados; devem começar a descrever a distribuição com base nas medidas de localização, amplitude e em gráficos que tenham elaborado com os dados. Além do cálculo da média, a associem à redistribuição equitativa; observem padrões ou tendências entre duas variáveis e interpretem os dados ao longo das diferentes representações numéricas e gráficas; e comecem a ganhar consciência das limitações das inferências, quando os dados são recolhidos na turma.

A um *nível de desenvolvimento intermédio*, os alunos devem saber analisar estatísticas sumárias de dados (medidas de localização, amplitude interquartil, desvio absoluto médio) e representações gráficas diversificadas (histogramas e diagramas de extremos e quartis, diagrama de dispersão e retas, entre outros), convenientes na comparação de dois ou mais conjuntos de dados; devem entender a média como um ponto de equilíbrio de uma distribuição de dados; devem interpretar graficamente o centro de gravidade de uma distribuição bivariada e o coeficiente de correlação; e devem modelar a associação linear através de uma reta e usá-la para estimar valores da variável dependente, conhecido um valor da independente. Os alunos aprendem a descrever dois conjuntos de dados atendendo às medidas e representações gráficas com respeito ao centro, dispersão e forma das distribuições e realizam interpretações à luz do contexto dos dados. Começam, também, a ter algum contacto com amostras não representativas e a distinguir estudos de observação de experiências, bem como a ganhar alguma sensibilidade relativamente ao tipo de resultados que cada um deles permite obter (associação ou causa-efeito).

A um *nível de desenvolvimento avançado*, os alunos devem ser capazes de identificar formas adequadas de organizar os dados e dar uso às medidas estatísticas de localização e de dispersão (nomeadamente ao desvio-padrão), fazendo leituras e interpretações adequadas a essas representações. Devem aprender a fazer afirmações e desenvolver as suas intuições sobre os dados à medida que o leque de representações numéricas e gráficas vai aumentando. Os alunos devem reconhecer associação entre duas variáveis qualitativas, e quando é que a associação entre duas variáveis quantitativas é razoavelmente linear; interpretar o coeficiente de correlação interligando-o com o centro de gravidade e entender quais são os dados que mais afetam o resultado; compreender o método dos mínimos quadrados, que permite encontrar a reta que melhor se ajusta aos dados; modelar a associação linear através da reta de regressão e interpretar os resultados à luz dos dados e problema (que devem ser do interesse dos alunos), entre outras capacidades.

2.3.2 A Estatística no Programa de Matemática A do ensino secundário

No programa de Matemática A do ensino secundário (ME, 2001), o tema da Estatística é apenas abordado no 10.º ano, para um conjunto de cerca de 15 aulas. Estabelece como conteúdos do ensino a desenvolver na unidade temática da Estatística os seguintes três tópicos gerais: a) Estatística - generalidades; b) Organização e interpretação de caracteres estatísticos (qualitativos e quantitativos); e c) Referência a distribuições bidimensionais. Além destes, existem temas transversais que atravessam o programa, independentemente da unidade temática e seus subtemas em estudo. São eles: Comunicação matemática; Aplicações e modelação matemática, História da matemática, Lógica e raciocínio matemático, Resolução de problemas e atividades investigativas, e Tecnologia e Matemática (ME, 2001).

Apresenta-se, de uma forma sumária, os conteúdos ao longo dos três tópicos gerais mencionados (adaptado de ME, 2001, pp. 29-31):

Estatística e Generalidades

- Objetivo da Estatística, breve nota histórica da sua evolução e utilidade na vida moderna. Clarificar com exemplos de fenómenos que podem ser objeto de um estudo estatístico.
- Recenseamento e sondagem. Noção de população e de amostra. Conceito de amostragem e reconhecimento do seu papel nas conclusões estatísticas. Distinção entre recenseamento e sondagem e que conclusões permitem retirar sobre a população. Noções intuitivas acerca da escolha de amostras (não enviesada, aleatoriedade e representatividade).
- Distinção entre Estatística descritiva e Estatística inferencial.

Organização e interpretação de carateres estatísticos

- Dados qualitativos representados em gráficos circulares, pictogramas e gráficos de barras e análise gráfica. Moda.
- Dados quantitativos (variável discreta/contínua), dados agrupados em classes.

Função cumulativa. Tabelas de frequências. Medidas de localização. Medidas de dispersão. Discussão em torno das limitações dessas medidas. Diagrama de extremos e quartis.

Referência a distribuições bidimensionais (gráfica e intuitiva)

- Relação bivariada; Diagrama de dispersão; associação linear, dependência estatística; ideia intuitiva de correlação (exemplos de gráficos com correlação positiva, negativa ou nula). - Coeficiente de correlação linear; centro de gravidade e sua interpretação. - Reta de regressão, ideia intuitiva, interpretação e limitações. Modelo linear e sua utilização na previsão de eventos.

Das indicações metodológicas especificamente dirigidas aos diferentes tópicos do tema da Estatística, destacam-se, sobretudo, que os professores devem dar atenção: i) ao papel relevante da Estatística relativamente às outras áreas do conhecimento; ii) à distinção entre Estatística descritiva e Estatística indutiva ou inferencial e realçar a importância do planeamento cuidado de um estudo estatístico; iii) aos cuidados no modo de aplicação dos instrumentos de redução de dados, nomeadamente, quais os mais ou menos adequados às diferentes situações e; iv) à abordagem gráfica e intuitiva no estudo dos dados bivariados. Ou seja, a partir da representação gráfica de dados bivariados, pode-se tentar averiguar a associação linear entre as duas variáveis em estudo. Caso esta exista, será possível chegar-se a um modelo matemático que permite obter uma estimativa de uma variável em estudo, sempre que é conhecido o valor da outra variável.

O programa de Matemática A do ensino secundário mais recente (MEC, 2014), no campo da Estatística, inclui apenas o tema *Caraterísticas amostrais* no 10.º ano, sugerindo que este seja desenvolvido em 18 aulas. De uma forma sumária apresenta-se os conteúdos deste tema do 10.º ano (adaptado de MEC, 2014, p. 14):

Caraterísticas amostrais

- Sinal de somatório e algumas das suas regras operatórias (uso das propriedades associativa e comutativa generalizada da adição e distributiva generalizada da multiplicação); - Interpretar e usar a variável estatística quantitativa; população e amostra; média (como fórmula e como centro de gravidade), variância e desvio-padrão de uma amostra; algumas propriedades básicas da média, variância e desvio-padrão de uma amostra. - Percentil de ordem k, propriedades do percentil de ordem k. Resolução de problemas envolvendo: média e desvio-padrão de uma amostra; percentis de uma amostra.

No tema Estatística do atual programa de Matemática A do ensino secundário (MEC, 2014), é introduzido o sinal de somatório e abordam-se algumas das suas regras operatórias. A noção de quartil do programa anterior (ME, 2001) é alargada com a noção de percentil de ordem *k* de uma amostra. Ao nível do 10.º ano, o programa (MEC, 2014) parece enfatizar especialmente o desenvolvimento de algumas propriedades matemáticas de algumas noções (tais como transformar o numerador da fórmula da variância $\sum_{i=1}^{n}(x_i - \bar{x})^2$ numa soma de três ou dois somatórios; determinar os resultados da variância/desvio-padrão, quando os dados quantitativos da amostra são todos multiplicados ou somados por um número real *a* e comparar esses resultados com os respetivos da amostra original), bem como a aplicação de cálculos e procedimentos em torno de todos os tópicos abrangidos. De uma maneira geral, o programa atual de Matemática A do ensino secundário (MEC, 2014) não valoriza as vertentes descritiva e inferencial da Estatística nem o modo como essas duas vertentes se complementam.

A organização dos dados e sua apresentação através de representações gráficas e numéricas, a utilização de um conjunto de noções estatísticas (quando considerados dados univariados: medidas de localização e de dispersão; quando considerados dados bivariados: associação linear, reta de regressão, coeficiente de correlação, centro de gravidade e modelação) e a indicação do potencial da tecnologia na Estatística são aspetos comuns entre o documento português (ME, 2001) e os documentos americanos (Franklin et al., 2007; NCTM, 2007). A atenção à variabilidade e ao processo investigativo não emerge de um modo explícito nos programas de Matemática A do ensino secundário no tema da Estatística (ME, 2001; MEC, 2014), enquanto, no caso das orientações curriculares internacionais referidas, esses assuntos são aspetos centrais e condutores do ensino da Estatística. Adicionalmente, a recolha de dados, atenção ao contexto dos dados, amostragem e inferência, associação e modelação, e o entendimento das probabilidades como uma ferramenta para a Estatística, nomeadamente, para aceder-se à Estatística inferencial, são assuntos que são mencionados de uma forma aprofundada nos documentos americanos, enquanto nos portugueses são praticamente inexistentes ou mencionados de uma forma bastante superficial. Além disso, os referidos documentos americanos expressam a relevância de se desenvolver a literacia, raciocínio e pensamento estatísticos dos alunos do ensino secundário de acordo com as definições apresentadas no trabalho de Garfield e Ben-Zvi (2007), ou seja, tendo em mente que o pensamento estatístico é o patamar mais elevado a alcançar e que todos os alunos que terminam o ensino secundário devem saber usar um sólido raciocínio estatístico. Estes termos não surgem em nenhum dos dois últimos programas de Matemática A do ensino secundário do nosso país.

CAPÍTULO 3

O CONHECIMENTO DIDÁTICO DO PROFESSOR EM ESTATÍSTICA

Este capítulo, tal como o anterior, insere-se no enquadramento teórico do estudo. É especialmente dedicado ao conhecimento didático do professor que ensina Estatística. Abrange quatro secções principais, estruturadas pela ordem seguinte: a natureza do conhecimento profissional do professor de Matemática; modelos relativos ao conhecimento didático do professor de Matemática, o conhecimento estatístico para ensinar e o conhecimento didático do professor em Estatística

Com essas secções mostra-se que a noção do conhecimento didático do professor que ensina Estatística desenvolveu-se fundamentalmente a partir do trabalho realizado na área do conhecimento profissional do professor de Matemática. Reportam-se modelos teóricos proeminentes da área da Educação Matemática (o modelo do *Pedagogical Content Knowledge* de Shulman (1986), o modelo do *conhecimento matemático para ensinar* de Ball (Ball et al., 2005, 2008), o modelo do *quarteto do conhecimento* de Rowland (Rowland, Huckstep & Thwaites, 2005) e o modelo do *conhecimento didático* do professor de Matemática de Ponte e Oliveira (2002) e de Canavarro (2003)) e da Educação Estatística (os modelos do *conhecimento estatístico para ensinar* de Groth (2007, 2013) e de Burgess (2006, 2008)). Incluem-se adicionalmente estudos realizados no campo da Estatística com base nesses modelos e também se procura fundamentar, com alguma discussão, o modelo adotado no estudo relativo ao *conhecimento didático do professor em Estatística*.

3.1 A natureza do conhecimento profissional do professor de Matemática

A investigação sobre o professor e seu ensino não é nova, mas ganhou sobretudo um novo fôlego nos anos 80 com as novas perspetivas trazidas por autores tais como Elbaz, Schön e Shulman que impulsionaram, em grande medida, a investigação na área (Ponte & Chapman, 2006), tanto do ponto de vista teórico como da produção de estudos empíricos (Chapman, 2013). Segundo Ponte e Oliveira (2002) e Ponte (2008, 2012), o conhecimento profissional do professor de Matemática é complexo, envolvendo o conhecimento da prática e de outros papéis assumidos pelo professor relacionados com a profissão. Apoia-se também no conhecimento de natureza teórica (sobre Matemática e educação), e no conhecimento de natureza social e experiencial (alunos, dinâmica da aula, comunidade escolar e profissional, formação, etc.) e inclui a visão que o professor possui sobre o seu próprio desenvolvimento profissional.

Num ensaio teórico sobre o conhecimento do professor, realizado por Guimarães (2008), são confrontadas as perspetivas de Schön, Elbaz e Clandinin que iniciaram nos anos 80 e 90 investigação nessa área e que contribuíram para o seu desenvolvimento. Guimarães (2008) indica que as perspetivas destes autores fazem emergir a valorização do papel do professor e o reconhecimento da importância do seu pensamento. O autor identifica três caraterísticas principais do conhecimento do professor: *experiencial*, uma vez que "tem origem na prática que os professores empreendem e que se desenvolve e manifesta nessa prática"; *situacional*, devido a ser "orientado para as situações que o professor enfrenta e, de certa maneira, indissociável dessas situações"; e *pessoal*, dado que "todo o conhecimento tem por base uma perspetiva própria de cada pessoa, estando impregnado pelas conceções, valores e propósitos que constituem aquela perspetiva" (p. 836). Neste ensaio, Guimarães (2008) conclui que o conhecimento profissional do professor diz respeito a um conhecimento específico, orientado para situações da sua prática "e que lhe permite interpretá-las, agir sobre elas e apreciar o resultado da sua atuação" (idem).

Segundo Munby, Russell e Martin (2001), capturar a natureza do conhecimento do professor passa por considerar-se a relação entre este conhecimento e a prática. De acordo com Ponte e Oliveira (2002) e Canavarro (2003), o conhecimento profissional do professor inclui um domínio que está intimamente ligado com a prática letiva, o *conhecimento didático*, que se trata de um conhecimento essencialmente dirigido para a ação. Este conhecimento não pode ser considerado como conhecimento declarativo ou formal (Ponte, 2008). Adicionalmente, Ponte e Santos (1998) referem-no como um conhecimento dinâmico que se refaz constantemente em função das experiências e das situações de prática escolar que o professor vai encontrando. Canavarro (2003) menciona ainda que é na ação que este conhecimento transparece, "pois o seu caráter implícito e tácito nem sempre o torna traduzível de forma proposicional" (p. 62).

A noção de conhecimento didático teve uma forte influência da noção de *Pedagogical content knowledge* (PCK) de Shulman (1986); ainda assim, o conhecimento didático (Ponte & Oliveira, 2002) surgiu desde cedo associado à ideia de que se trata de um conhecimento complexo e que não tem o propósito de determinar um conjunto de capacidades básicas ou de respostas possíveis para problemas específicos do contexto escolar.

Na literatura em Educação Matemática, o PCK é frequentemente traduzido como um domínio do conhecimento profissional que tem a particularidade de combinar capacidades do professor ao nível do conteúdo matemático e do seu ensino (por exemplo, Ball et al., 2008; Depaepe, Verschaffel & Kelchtermans, 2013). Depaepe et al. (2015) defendem que a investigação na área do conhecimento do professor será mais profícua quanto mais próxima estiver das práticas letivas e contemplar na análise deste conhecimento aspetos afetivos que possam influenciar a atuação do professor, além dos aspetos cognitivos.

3.2 Modelos relativos ao conhecimento didático do professor de Matemática

3.2.1 O modelo do PCK de Shulman

Shulman (1986), ao notar que a maioria dos trabalhos de investigação sobre o ensino se centrava em aspetos gerais (gestão de sala de aula, planificação de aulas, distribuição do tempo ou organização de atividades), chamou a atenção para que os estudos sobre o conhecimento do professor problematizassem o papel do conteúdo no ensino. Este autor designou este problema por "padrão omisso da investigação no ensino", referindo: "O que faz falta são questões sobre o conteúdo das lições ensinadas, as questões feitas e as explicações dadas" (Shulman, 1986, p. 8).

Os trabalhos de investigação de Shulman conduziram ao desenvolvimento de uma categorização do conhecimento profissional do professor. Shulman (1986) apresentou três categorias distintas deste conhecimento: (i) Conhecimento do conteúdo; (ii) Conhecimento pedagógico do conteúdo (*PCK-Pedagogical content knowledge)*; e (iii) Conhecimento curricular. De seguida expõem-se os elementos e aspetos mais importantes de cada uma destas três categorias.

Para Shulman (1986), o *conhecimento do conteúdo* abrange as estruturas substantivas e as estruturas sintáticas do conhecimento (citando Schwab, 1978). As primeiras referem-se aos modos como os princípios e conceitos de uma disciplina estão organizados, as segundas referem-se a um conjunto de regras estabelecidas no âmbito de uma disciplina. Ainda a este respeito, Shulman menciona que não basta aos professores de uma área disciplinar dominar estas estruturas; além de saber avaliar a veracidade ou validade de um assunto relativamente a um determinado conteúdo, devem saber explicar o porquê da questão, bem como compreender a sua relevância e a relação deste conteúdo com outros. O autor destaca, assim, a importância de os professores possuírem uma compreensão fundamentada sobre quais deverão ser os tópicos centrais e os mais periféricos da disciplina. O conhecimento do conteúdo diz respeito à quantidade e organização do conhecimento do professor na sua mente. Este conhecimento envolve ir além dos conhecimentos dos factos e conceitos de um domínio disciplinar.

Shulman (1986) salienta o *conhecimento pedagógico do conteúdo* (*PCK*) como um dos aspetos centrais a ser tomados em conta no estudo do ensino, por ser um conhecimento de conteúdo específico da atividade de ensinar. Este conhecimento engloba fazer-se uso de uma variedade de representações, exemplos, analogias, ilustrações, explicações e demonstrações dos conteúdos de um domínio disciplinar de modo a que os conteúdos escolares se tornem compreensíveis e acessíveis para os alunos. Este conhecimento abrange também uma compreensão dos aspetos que facilitam ou dificultam a aprendizagem de determinados conteúdos pelos alunos.

O *conhecimento curricular* abrange o conhecimento que o professor tem dos diferentes programas de um domínio disciplinar, dos recursos alternativos para o seu ensino e das vantagens/desvantagens do uso destes programas curriculares e respetivos materiais em certos contextos escolares. Shulman (1986) refere ainda dois aspetos principais deste conhecimento: o conhecimento horizontal do currículo e o conhecimento vertical do currículo. O primeiro diz respeito à capacidade do professor interligar os conteúdos da sua disciplina com os conteúdos trabalhados noutras disciplinas frequentadas pelos alunos. O segundo inclui a familiaridade com os conteúdos e respetivos materiais curriculares e com questões que possam surgir no seu ensino ao longo dos diferentes níveis escolares de uma área disciplinar.

Na sua perspetiva, a experiência que o professor vai adquirindo ao longo da sua carreira e os elementos provenientes da investigação em Educação (ensino e aprendizagem) devem contribuir para o desenvolvimento do seu *PCK* (Shulman, 1986). Este quadro teórico de Shulman (1986) teve um grande impacto na investigação sobre o conhecimento profissional do professor, tendo inspirado muitas investigações nesta área, em particular, no campo da Educação Matemática (Ponte & Chapman, 2006). Por exemplo, Oliveira e Ponte (1997) analisaram 76 artigos publicados em quatro anos consecutivos, de 1992 a 1995, provenientes de revistas e atas de conferências relevantes em Educação Matemática, nomeadamente, do *Journal for Research in Mathematics Education* (JRME), do *Educational Studies in Mathematics* (ESM) e da *Psychology of Mathematics Education* (PME). Este trabalho tinha como propósito traçar um quadro geral da investigação sobre a figura do professor de Matemática no que se refere a conceções, saberes e desenvolvimento profissional. Foram identificados cerca de 30 estudos que incidiam sobre os conhecimentos de base do professor que, segundo Oliveira e Ponte (1997), diziam respeito a "conhecimento sobre a Matemática, sobre o seu ensino e ainda o conhecimento mais diretamente ligado com a aprendizagem e o aluno" (p. 3). Estes mesmos trabalhos tiveram como quadro teórico de referência, de modo implícito ou explícito, o modelo de Shulman (1986). Segundo Oliveira e Ponte (1997), este modelo de Shulman, além de destacar a relevância dos conhecimentos de conteúdo e da pedagogia salienta um domínio especial que resulta da fusão desses dois conhecimentos, o conhecimento didático.

Registam-se também algumas investigações recentes que recorrem à noção de PCK de Shulman (1986) para estudos centrados no conhecimento do professor na área da Estatística, nomeadamente os realizados por Jane Watson e Rosemary Callingham da

Universidade de Tasmania (Austrália) (por exemplo, Watson & Callingham, 2013, que é descrito na secção 2.2.3 do capítulo 2).

3.2.2 O modelo do conhecimento matemático para ensinar

O modelo do PCK de Shulman inspirou o desenvolvimento de outros quadros teóricos sobre o conhecimento do professor de Matemática (Ball & Bass, 2005; Ball, Thames & Phelps, 2005; Rowland, Huckstep & Twaites, 2005). Principalmente, desde 2000, as noções de *conhecimento de Matemática* e *conhecimento pedagógico de conteúdo* (PCK) de Shulman têm passado por algum refinamento no desenvolvimento de um quadro teórico do *conhecimento matemático para ensinar* (Ball & Bass, 2005; Ball, Thames & Phelps, 2005, 2008).

Ball (2002) refere que, no início da sua carreira como professora do ensino básico, tomou a opção de voltar à universidade para frequentar disciplinas de Matemática, dado que, enquanto aluna do ensino superior, apenas tinha frequentado uma disciplina desta área. Esta experiência, segundo a autora, influenciou muito positivamente a sua prática profissional:

> A minha introdução no mundo das conjeturas e provas, de lemas e teoremas, fascinou-me e aliciou a minha imaginação como professora. (…) Comecei a puxar um pouco mais pelos meus alunos. Quando eles reparavam em padrões, ou experimentavam ideias, pedia-lhes para explicarem os seus raciocínios e para darem esclarecimentos das suas afirmações. (…) Quanto mais puxava os alunos para raciocinarem, mais eles pareciam compreender as ideias com as quais estávamos a trabalhar. (p. 2)

A autora denota ainda que, nos momentos em que lhe surgiam algumas dúvidas sobre aspetos matemáticos ou até ideias erradas, esta situação influenciava negativamente a sua capacidade para orientar tarefas e proporcionar uma discussão relevante em contexto de sala de aula. Deste modo, conclui que há uma ligação entre o ter melhorado a sua prática como professora de Matemática e o ter enriquecido o seu conhecimento matemático (Ball, 2002). Ball tem vindo desde então, juntamente com outros investigadores, a desenvolver trabalhos que ajudam a perceber a natureza do conhecimento matemático que parece contribuir para um ensino mais eficaz, designado por *conhecimento matemático para ensinar* (*mathematical knowledge for teaching)*, ou seja, trata-se do conhecimento matemático do professor que é especialmente útil no ensino da Matemática.

Esta autora e colaboradores consideram que o trabalho de Shulman (1986) foi fundamental para a investigação do conhecimento do professor ao atribuírem um papel de destaque ao conteúdo, no ensino, e ao reformularem o quadro teórico nesta área. Reconhecem também que o trabalho de Shulman contribuiu para definir o ensino como uma profissão ao focar-se na compreensão do conhecimento do conteúdo como um conhecimento técnico essencial para a atividade do professor (Ball et al., 2008).

Ball, Thames e Phelps (2008) definem o *conhecimento matemático para ensinar* como o "conhecimento matemático necessário para levar a cabo o trabalho de ensinar matemática aos alunos" (p. 395). Do ponto de vista concetual, este conhecimento deve ser entendido em termos mais gerais, ou seja, deve incluir capacidades, hábitos mentais e perceções. Além disso, é assumido pelos autores que a atividade de ensinar não se reduz ao mero uso de uma certa abordagem de ensino, mas envolve identificar tarefas essenciais para o ensino, bem como as exigências matemáticas de tais tarefas.

Ball e colaboradores procederam ao desenvolvimento concetual do *conhecimento matemático para ensinar*, com base no trabalho Shulman (1986) e através do desenvolvimento de investigações centradas na atividade de ensinar do professor tais como o *Mathematics Teaching and Learning to Teach Project* e o *Learning Mathematics for Teaching Project*. Esta linha de trabalho permitiu criar e definir subdomínios da noção de PCK de Shulman e expandir o conceito do *conhecimento matemático para ensinar* (ver figura 9). Detetou-se, em particular, um subdomínio classificado como conhecimento especializado do conteúdo que, apesar de se tratar de um conhecimento "essencial para um ensino eficaz" (Ball, Thames & Phelps, 2005, p. 390), não está incorporado no conhecimento pedagógico de conteúdo, mas no conhecimento de conteúdo. Relativamente a este conhecimento os autores acrescentam que

> o ensino pode requerer uma forma especializada de conhecimento disciplinar puro – puro porque não é misturado com conhecimento de alunos ou pedagogia e é então distinto do conhecimento pedagógico do conteúdo de Shulman e colaboradores e especializado porque não é necessário ou usado em contextos que não sejam o do ensino da matemática. Esta unicidade é aquilo que torna este conhecimento de conteúdo especial. (p. 396)

Ball et al. (2005) consideram que o *conhecimento especializado do conteúdo* é um conhecimento matemático específico do professor de Matemática, e é sobretudo usado em contextos relacionados com o ensino da disciplina.

O modelo do *conhecimento matemático para ensinar* integra então duas componentes principais – o *conhecimento do conteúdo* e o *conhecimento pedagógico do conteúdo* – e cada uma delas encontra-se dividida em três componentes. Assim, o *conhecimento do conteúdo* abrange *conhecimento comum do conteúdo*, *conhecimento especializado do conteúdo* e *conhecimento do horizonte*; o *conhecimento pedagógico de conteúdo (PCK)* abrange *conhecimento do conteúdo e do ensino*, *conhecimento do conteúdo e dos alunos* e ainda *conhecimento do conteúdo e do currículo*. O esquema seguinte (figura 9) ilustra a estrutura e as componentes do *conhecimento matemático para ensinar*, de acordo com Ball, Thames e Phelps (2008):

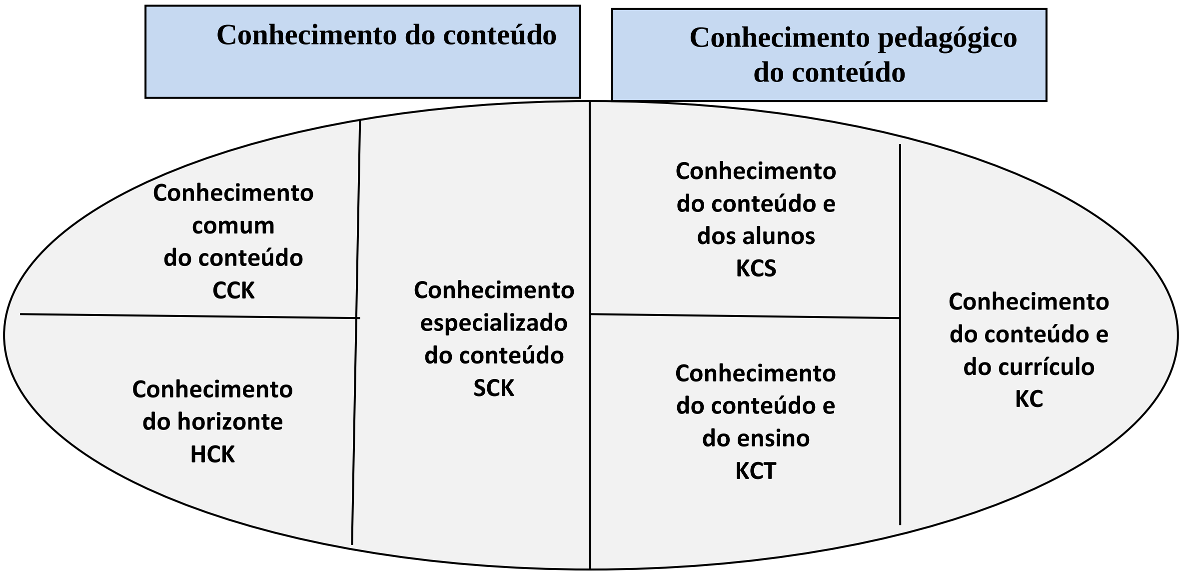


Figura 9: Domínios (e subdomínios) do conhecimento matemático para ensinar (adaptado de Ball, Thames & Phelps, 2008, p. 403).

Apresenta-se, em seguida, uma descrição dos domínios do *conhecimento matemático para ensinar* de acordo com Ball et al. (2008):

O *conhecimento comum do conteúdo* (*Common content knowledge – CCK*) é um conhecimento que combina conhecimento matemático e capacidades que podem ser usadas em vários contextos (por exemplo, por vendedores, enfermeiros, bancários, etc.). É, portanto, um conhecimento que não é exclusivo do ensino. Alguns exemplos deste conhecimento são a destreza no cálculo, resolver corretamente certos problemas matemáticos, reconhecer respostas erradas dos alunos, detetar definições incorretas em manuais escolares, realizar os trabalhos a propor aos alunos, etc.;

O *conhecimento do horizonte* do conteúdo (*Horizon Content knowledge – HCK*) diz respeito ao conhecimento do modo como os tópicos matemáticos se relacionam e se estendem ao longo do currículo dos diferentes níveis escolares;

O *conhecimento especializado do conteúdo* (*Specialized content knowledge – SCK*) é um conhecimento específico para o ensino que reúne conhecimento matemático e capacidades que não são habitualmente usadas fora do contexto do ensino. Contudo, este conhecimento diferencia-se do conhecimento de alunos ou pedagogia. De facto, o ensino exige um conhecimento disciplinar que está para além do que é ensinado aos alunos; no entanto, é essencial saber-se o modo como este conhecimento é produzido e está estruturado e as implicações destes aspetos no ensino.

Além disso, o professor deve fazer uso de um conhecimento matemático descompactado que permita que os alunos se apercebam e aprendam as características de um determinado conteúdo. Ao escolher, fazer e usar representações matemáticas de modo eficaz, reconhecendo as mais ou menos úteis na explicação de uma definição ou de um determinado conteúdo; ao explicar e justificar ideias matemáticas, ao dar explicações explícitas sobre questões de linguagem matemática e ao verificar se um novo método usado por um aluno faz sentido do ponto de vista matemático, os professores utilizam a Matemática de modo especializado (Ball et al., 2008);

O *conhecimento do conteúdo e dos alunos* (*Knowledge of content and students–KCS*), é um conhecimento que diz respeito à capacidade do professor antecipar erros e escutar e interpretar raciocínios dos alunos na sua própria linguagem, bem como de prever o modo como estes lidam com determinadas tarefas, ou seja, avaliar em quais poderão ter mais facilidade ou dificuldade e em que aspetos específicos, e quais serão mais ou menos interessantes para eles. Estas tarefas requerem uma interação entre compreensão matemática específica e familiaridade com os alunos e com o seu pensamento matemático;

O *conhecimento do conteúdo e do ensino* (*Knowledge of content and teaching – KCT*) é um conhecimento que diz respeito à capacidade do professor planear a sequência dos conteúdos, reconhecer as vantagens e desvantagens do uso de diferentes representações no ensino de um tópico específico, avaliar e selecionar quais os métodos e procedimentos mais úteis no ensino dos diferentes tópicos, tomar decisões sobre as contribuições que deve solicitar aos alunos. Estas tarefas requerem uma interação entre compreensão matemática específica e compreensão de questões pedagógicas que podem afetar a aprendizagem dos alunos. O *conhecimento do conteúdo e do currículo* (*Knowledge of content and curriculum –KCC*) é um conhecimento que diz respeito ao conhecimento curricular, tal como Shulman (1986) o descreve (Ball et al., 2008).

Ball et al. (2005, 2008) reconhecem que há situações da prática em que é difícil distinguir o *conhecimento comum* do *conhecimento especializado* e, noutras situações, o *conhecimento especializado* do *conhecimento do conteúdo e alunos*. Por exemplo, o conhecimento acerca da fração que corresponde a uma área sombreada específica de dois círculos geometricamente iguais é *conhecimento comum* ou *conhecimento especializado*? Os autores consideram que este exemplo poderá estar mais relacionado com o *conhecimento especializado,* embora admitam que possa haver pessoas que no seu trabalho se baseiem neste conhecimento pormenorizado e, assim, tratar-se-ia de um *conhecimento comum*. Outro exemplo que os autores apresentam permite também perceber a dificuldade em distinguir, por vezes, o *conhecimento especializado* do *conhecimento de conteúdo e alunos*: ao analisar um erro cometido por um aluno, o professor usa um *conhecimento especializado* se analisar o erro matematicamente ou pode fazer uso de um *conhecimento de conteúdo e alunos,* se já tiver contactado com alunos que fizeram o mesmo tipo de erro. Para os autores, todas as categorias que desenvolveram precisam de revisão e de aperfeiçoamento e reforçam a importância da continuação deste trabalho em prol do desenvolvimento profissional do professor (Ball et al., 2005, 2008).

De acordo com este quadro teórico, o professor deve dominar e desenvolver um conhecimento matemático descompactado que permita que os aspetos importantes dos conteúdos a ensinar sobressaiam e que as ideias matemáticas se tornem mais acessíveis aos alunos. Este conhecimento descompactado deve estar presente na atividade do professor, em particular, quando utiliza o *conhecimento do conteúdo* (conhecimento comum e especializado) e o *conhecimento pedagógico do conteúdo* (conhecimento matemático e dos alunos e conhecimento matemático e do ensino). Já os alunos, por seu turno, nas suas aprendizagens devem desenvolver uma certa espontaneidade com o conhecimento matemático através de ideias intuitivas, propriedades e noções formais (Ball et al., 2005, 2008).

Algumas componentes deste modelo têm surgido complementadas com outras noções teóricas, nomeadamente, relacionadas com Estatística e contempladas em investigações que visam estudar o *conhecimento do professor que ensina Estatística* em diferentes níveis de ensino (e.g., Burgess, 2006, 2009, 2011; Groth, 2007, 2013; Oliveira & Henriques, 2014).

O modelo do *conhecimento matemático para ensinar* (Ball, Thames & Phelps, 2008) foi usado por Ribeiro, Carrillo e Monteiro (2009) para analisar o conhecimento profissional de uma professora de Matemática do primeiro ciclo a partir da sua prática, nomeadamente, quando esta pretendia que os alunos realizassem inferências a partir de um pictograma específico (cujos dados tinham sido recolhidos juntos dos alunos numa aula). Os autores concluíram que quatro das componentes do *modelo do conhecimento para ensinar* foram evidenciadas na análise efetuada, a saber, do *conhecimento do conteúdo*, o conhecimento *comum* e o *especializado*, do *conhecimento pedagógico do conteúdo (PCK)*, o conhecimento de *conteúdo e do ensino* e o de *conteúdo e dos alunos*. Os autores identificaram algumas dificuldades e lacunas da professora ao nível do *conhecimento matemático* e do *conhecimento pedagógico* e consideraram que estudos desta natureza podem auxiliar e orientar o trabalho a realizar em termos de formação inicial e contínua na área da Estatística e do seu ensino.

Para estudar o conhecimento profissional do professor de Matemática do ensino secundário relativamente ao ensino da variabilidade, na Estatística, González (2013) usou o modelo do *conhecimento matemático para ensinar* (Ball et al., 2008) com as suas seis componentes, e articulou-o com uma vertente afetiva, ou seja, considerando as crenças dos professores relativamente à Estatística escolar e à aprendizagem. O autor afirma que essa vertente afetiva não podia ser ignorada no seu estudo visto ser um fator que condiciona a prática.

González (2013) apresenta alguns resultados preliminares sobre o conhecimento profissional do professor com base em respostas abertas a duas alíneas de uma tarefa que foram respondidas anonimamente através de um questionário. Para este trabalho foram consideradas as respostas individuais de um grupo de quatro professores do ensino secundário de uma escola secundária japonesa. Com estes dados, a primeira alínea iria permitir analisar o *conhecimento comum* (que o autor associou à literacia estatística) dado que incluía duas distribuições gráficas e pedia a indicação e justificação da que apresentava maior variabilidade. Já a segunda alínea iria permitir analisar o *conhecimento de conteúdo e do ensino* e a vertente afetiva (referida acima), através de um pedido de elaboração de um plano de aula que abordasse a variabilidade, tomando a alínea inicial como exemplo. O autor refere que, na primeira alínea, houve um professor que usou apenas uma estratégia na comparação das distribuições (diferença entre a média e um dado extremal), enquanto outros dois usaram mais estratégias que incluíam o uso e cálculo de medidas variabilidade diferentes (intervalo interquartil, amplitude interquartil, variância, desvio à média). Na segunda alínea, todos os professores indicaram alguns aspetos relevantes de como lidar com a variabilidade no ensino; ainda assim, nunca assinalaram, por exemplo, a necessidade de comparar as distribuições com outras representações gráficas e houve, inclusivamente, um professor que não referiu a importância de considerar situações reais no estudo da variabilidade. Quanto à vertente afetiva, González (2013) mencionou haver dois professores que pareciam perspetivar a aprendizagem da Estatística, no ensino secundário, centrada no professor enquanto os outros dois, centrada no trabalho do aluno. O autor considerou que, de uma maneira geral, os participantes deste estudo pareciam revelar não conhecer abordagens diversificadas no ensino da variabilidade.

3.2.3 O modelo do quarteto do conhecimento

O *quarteto do conhecimento* (*knowledge quartet*) é um modelo que se inspirou no trabalho de Shulman (1986) e foi desenvolvido para identificar, descrever e analisar o conhecimento do conteúdo matemático evidenciado no ensino, com o propósito de disponibilizar um quadro teórico útil para a reflexão e discussão de aulas, como base para o desenvolvimento profissional do futuro professor ou de professores no ativo (Rowland, Huckstep & Twaites, 2005; Rowland, Turner & Twaites, 2008; Turner &

Rowland, 2011). O construto de reflexão que lhe está associado tem origem nos trabalhos de Schön (1983).

O modelo do *quarteto do conhecimento* (Rowland et al., 2005, 2008) tem uma base empírica iniciada em 2002/2003 que resultou da análise de um total de 24 aulas, em que cada um dos 12 professores participantes lecionou duas aulas (professores em formação e futuros professores do ensino básico). Desta análise foram identificados 20 códigos referentes a situações diversificadas (Quadro 2), que posteriormente foram agrupadas em quatro categorias amplas e, consequentemente, deram origem ao *quarteto do conhecimento*: *Fundação* (*foundation*), *Transformação*, (*transformation*), *Conexão* (*connection*) e *Contingência* (*contingency*). A primeira (*fundação*) inclui o conhecimento matemático adquirido pelo professor durante a sua formação, na escola e na universidade. Engloba ainda crenças e perspetivas do professor relativamente ao ensino e aprendizagem. A segunda (*transformação*) é influenciada pela ideia de que o professor transforma o conhecimento matemático de forma a torná-lo compreensível aos alunos através de “analogias, exemplos, explicações e demonstrações” (Rowland et al., 2008, p. 3). A terceira categoria (*conexão*) diz respeito à coerência da planificação e do ensino ao longo das aulas, o que inclui a forma como é concebida a sequência dos conteúdos no ensino e também a noção das exigências cognitivas ao longo dos diferentes tópicos e tarefas propostas. A última componente (*contingência*) diz respeito ao modo como o professor lida com imprevistos que surjam no decorrer das aulas (Rowland et al., 2008).

Quadro 2: Códigos associados a cada dimensão do *Quarteto do conhecimento* (Rowland et al., 2008)

Quarteto do conhecimento	Códigos descritores
Fundação	Consciência dos objetivos, adesão ao manual, foco nos procedimentos, identificação de erros, apresentação direta dos conhecimentos, fundamentação teórica da pedagogia, utilização de terminologia matemática.
Transformação	Escolha de exemplos, escolha e uso de representações, uso de materiais instrucionais e demonstração pelo professor para explicar um procedimento.
Conexão	Antecipação da complexidade, decisões sobre a sequenciação, conexões entre procedimentos, ligações entre conceitos, reconhecimento sobre a adequação concetual.
Contingência	Desvio da planificação, resposta às ideias dos alunos, exploração das oportunidades, compreensão do professor no decorrer do processo de ensino.

Rowland (2013) refere que o *quarteto do conhecimento* pode ser visto como uma ferramenta que permite ao professor fazer uma análise reflexiva sobre o conteúdo matemático usado na sua prática. O autor considera que o *conhecimento matemático para ensinar* (Ball et al., 2008) e o *quarteto do conhecimento*, apesar de terem em comum emergirem a partir da prática, são dois modelos com propósitos diferentes. Na sua perspetiva, no *quarteto do conhecimento* a distinção das suas componentes é importante; contudo, considera que é ainda mais relevante a classificação de situações nas quais o conhecimento matemático emerge no ensino; no *conhecimento matemático para ensinar* há o objetivo de descortinar as noções de conhecimento matemático e de conhecimento pedagógico de conteúdo. Assim sendo, o autor considera que os dois modelos podem ser vistos e usados como modelos complementares.

Segundo Rowland (2013), há vários exemplos na literatura que envolvem matemática do ensino básico e que percorrem as quatro componentes do modelo do *quarteto do conhecimento*. Este modelo continua em desenvolvimento, passando a integrar, mais recentemente, uma coleção de exemplos de episódios de aulas para cada um dos códigos descritores de cada componente, envolvendo conteúdo matemático do ensino básico e do ensino secundário. Esta foi uma solução encontrada pelo autor e seus colegas de equipa contra o risco de poder surgir, nomeadamente, na investigação, um leque variado de interpretações para cada código descrito.

Num estudo longitudinal, de cerca de quatro anos, sobre o conhecimento do professor de Matemática, que tinha como propósito usar o modelo do *quarteto do conhecimento* como ferramenta para identificar e fomentar os conhecimentos de *conteúdo* e *pedagógico* do professor, foram constituídos quatro estudos de caso de professores de Matemática do ensino básico em início de carreira (Turner, 2008; Turner & Rowland, 2011). Esses participantes tiveram oportunidade de se familiarizar com o modelo do *quarteto do conhecimento*, assumido no estudo, durante o primeiro ano da investigação e de o usar de modo mais independente posteriormente, embora sem nunca terem deixado de discutir a aplicação do modelo nas reuniões de grupo e com o investigador (Turner, 2008). Esses quatro casos foram formados através de dados oriundos de fontes diversificadas: observação de aulas, entrevistas de reflexão sobre a prática pós-aula, relatórios reflexivos de cada participante de análise de episódios vídeo gravados das suas aulas através do modelo, entrevistas em grupo e entrevistas individuais (Turner, 2008; Turner & Rowland, 2011). Os participantes deram valor ao facto de receberem *feedback* relativamente ao conteúdo matemático que desenvolveram nas suas aulas e indicaram a utilidade do modelo do *quarteto do conhecimento* na planificação e na avaliação do ensino. Ainda, nesta investigação, Turner e Rowland (2011) concluíram que este modelo é uma ferramenta útil para o desenvolvimento do conhecimento do conteúdo matemático do professor através da reflexão sobre a sua prática. O uso do modelo parece ter ajudado os participantes a desenvolver a capacidade de não se focarem sobretudo no ensino dos procedimentos, mas também passarem a valorizar o entendimento concetual dos assuntos a ensinar. O modelo parece ter contribuído para que os professores em início de carreira desenvolvessem o seu PCK e identificassem e melhorassem aspetos do seu conhecimento matemático.

3.2.4 O modelo do conhecimento didático do professor de Matemática

Vários estudos indicam que os professores de diferentes níveis de ensino devem saber mais do que aquilo que ensinam (e.g., Ponte, 2012; Watson & Harel, 2013). Deter um bom conhecimento do conteúdo é importante, mas não é suficiente para um ensino eficiente (Depaepe et al., 2013, 2015; Ma, 1999; Petrou & Goulding, 2011; Ponte, 2014; Rowland et al., 2008). De acordo com Ponte e Oliveira (2002) e Ponte (2012), o conhecimento profissional do professor de Matemática desdobra-se por diversos domínios, nomeadamente, o conhecimento na ação relativo à prática letiva (o conhecimento didático), à prática não letiva, à profissão e ao desenvolvimento profissional. O conhecimento profissional do professor tem uma natureza multidimensional e as categorias que o integram estão obviamente interrelacionadas (Chapman, 2013; Ponte & Oliveira, 2002; Petrou & Goulding, 2011; Ponte, 2012). Assim, o conhecimento do professor de Matemática é muito mais abrangente do que o *conhecimento matemático para ensinar* de Ball e colegas da universidade de Michigan, que está organizado tendo por base o conteúdo matemático e duas categorias principais o *conhecimento do conteúdo* e o *PCK* (Beswick, Callingham & Watson, 2012).

Depaepe et al. (2013) e Ponte e Chapman (2006) referem que a noção de *PCK* de Shulman (1986) tem sido alvo de críticas, inclusive do próprio Shulman, por não incluir certos aspetos que são relevantes no ensino tais como o nível de ação do professor, a motivação e a paixão na atividade do professor. Para além disso, o *PCK* é uma noção que toma o indivíduo como unidade de análise sem dar atenção aos professores como um grupo profissional; por conseguinte, esta noção deve ser expandida, voltada para a ação ou apoiada na prática, passando a incluir o currículo, os alunos e a comunidade (Ponte & Chapman, 2006). Também Petrou e Goulding (2011) apontam que, no trabalho de Shulman, faltou a identificação das interações entre as diferentes categorias do conhecimento profissional do professor.

Para Ponte e Oliveira (2002) e Ponte (2012), este *conhecimento didático do professor* organiza-se em quatro grandes domínios (figura 10), que não podem ser considerados como estanques uma vez que interagem uns com os outros e estão presentes na atividade do professor quando ensina Matemática. De seguida, descreve-se cada um desses domínios: (i) conhecimento da Matemática; (ii) conhecimento do aluno e da aprendizagem; (iii) conhecimento do currículo; e (iv) conhecimento do ensino.

O *conhecimento da Matemática* inclui o conhecimento de conceitos matemáticos, procedimentos e representações a ensinar, e a interpretação que o professor faz da disciplina escolar, incluindo conexões internas entre diversos tópicos. Este conhecimento abarca também perceber com que profundidade cada assunto deve ser tratado. O *conhecimento do aluno e da aprendizagem* engloba uma compreensão sobre interesses, necessidades e dificuldades mais frequentes dos alunos, bem como de aspetos culturais e sociais que possam interferir positiva ou negativamente no seu desempenho escolar. Integra também o conhecimento das estratégias e dos processos de aprendizagem dos alunos. O *conhecimento do currículo* diz respeito às grandes finalidades e objetivos da disciplina e à articulação vertical e horizontal do currículo, bem como ao conhecimento de recursos diversificados e de formas de avaliação a utilizar. A leitura que os professores fazem dos programas escolares e aquilo que pretendem alcançar influencia a sua atuação na prática. O *conhecimento do ensino*, no que se refere à preparação, condução e avaliação da prática letiva, diz respeito a tudo o que tem a ver com a condução efetiva do ensino de Matemática, incluindo a preparação antes da lecionação, decisões em termos de tarefas, modos de organização do trabalho na aula, as próprias interações na aula, avaliação da aprendizagem dos alunos e reflexões do professor após a lecionação. Este conhecimento é focado na prática letiva e, na perspetiva de Ponte (2012), deve ser considerado o núcleo fundamental do conhecimento didático.

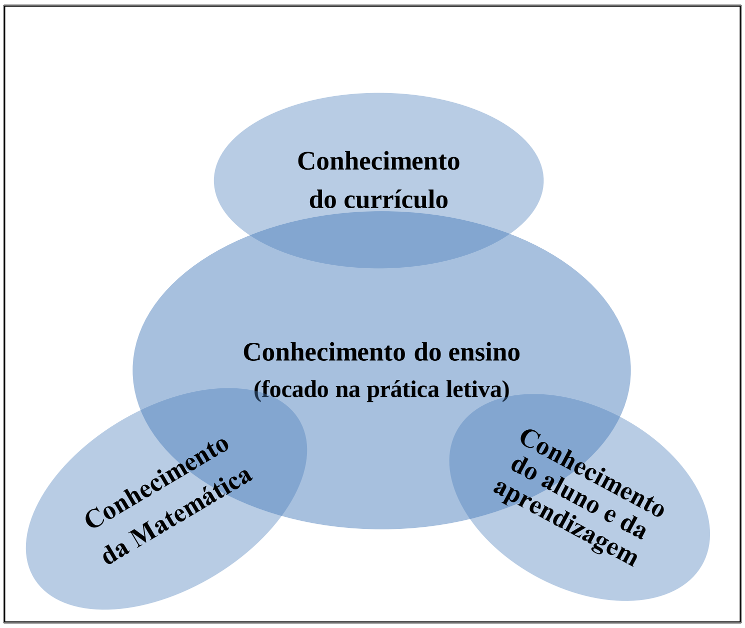


Figura 10: Dimensões do conhecimento didático (adaptado de Ponte, 2012).

Segundo Ponte (2012), este modelo do conhecimento didático tem a particularidade de, por um lado, assumir que tem uma dimensão central – o conhecimento do ensino – que se serve das outras três dimensões do modelo; e por outro lado, considerar que as suas quatro dimensões não estão separadas, visto qualquer uma delas estar presente na atividade do professor que ensina Matemática. O autor reconhece que essas duas caraterísticas do modelo distinguem-no de outros supostamente idênticos.

No contexto nacional, a noção de *conhecimento didático* surge ligada ao exercício profissional (Ponte, 2008) e está muito presente na investigação em Educação Matemática, dos últimos cerca de 25 anos, em Portugal (e.g., Canavarro, 2003; Duarte, 2011; Guimarães, 1999; Oliveira, 1998; Ponte & Santos, 1998; Ribeiro et al., 2009). Ponte (2012) refere ainda que a investigação nessa área inspirada neste modelo do conhecimento profissional do professor de Matemática tem contribuído “com elementos significativos para o compreender, nomeadamente, no quadro das exigências colocadas pelas orientações curriculares e pelas expectativas da sociedade em relação à escola” (p. 89).

De acordo com Batanero e Godino (2005) e Godino et al. (2008), o conhecimento profissional do professor que ensina Estatística, inclui, além de conhecimento de

Matemática e de Estatística, o *conhecimento didático* que abarca os seguintes cinco aspetos complementares: (1) epistemologia – reflexão epistemológica sobre os significados dos conceitos a ensinar; (2) cognição – capacidade de prever dificuldades de aprendizagem, erros, obstáculos e estratégias dos alunos; (3) recursos e metodologias de ensino – experiência com bons exemplos de situações de ensino, ferramentas didáticas e ainda capacidade crítica para analisar manuais, documentos curriculares e para adaptar o conhecimento estatístico a diferentes níveis de ensino e mostrar o caráter interdisciplinar da Estatística; (4) afeto – capacidade de incorporar os interesses dos alunos e ter em conta as suas atitudes e crenças; e (5) interação – capacidade de criar boa comunicação na aula, avaliar o trabalho desenvolvido no ensino e promover uma interação que favoreça a supressão das dificuldades que emergiram no ensino (Godino et al., 2008).

Estes cinco aspetos estão em consonância e interrelacionam-se com as dimensões do *conhecimento didático do professor* de Ponte e Oliveira (2002) e de Ponte (2012). Ou seja, a *epistemologia* relaciona-se com o *conhecimento da Matemática*; a *cognição* relaciona-se essencialmente com o *conhecimento do aluno e da aprendizagem*, mas também pode ser evidenciada no *conhecimento do ensino*; os *recursos e metodologias de ensino* relacionam-se com o *conhecimento do ensino* e com o *conhecimento do currículo*; o *afeto* e a *interação* relacionam-se com o *conhecimento do aluno* e com o *conhecimento do ensino*.

Adicionalmente, as dimensões do *conhecimento didático do professor* de Ponte e Oliveira (2002) e de Ponte (2012) também se harmonizam com os quatro domínios do modelo do quarteto do conhecimento. Ou seja, *Fundação* tem uma certa correspondência com o *conhecimento da Matemática* e com o *conhecimento do currículo*; *Transformação, Conexão* e *Contingência* são domínios que são mobilizados na prática e, assim sendo, emergem no *conhecimento do ensino*; além disso esses três domínios têm alguns pontos comuns com o *conhecimento do aluno e da aprendizagem*, quando se tem em conta o conhecimento matemático mobilizado pelo professor para fazer com que os alunos aprendam. Os quatro domínios do quarteto têm também alguns aspetos comuns com o *conhecimento do currículo*.

O modelo do *conhecimento didático do professor* de Ponte e Oliveira (2002) e de Ponte (2012) é o usado na presente investigação; contudo, foi adaptado, sobretudo cruzado com outros elementos, para atender à especificidade da Estatística e do seu ensino de acordo com diversos autores (e.g., Burrill & Biehler, 2011; Franklin et al., 2007; Rossman et al., 2006; Scheaffer, 2006; Shaughnessy, 2007), como será esclarecido, mais à frente, neste capítulo.

3.3 O conhecimento estatístico para ensinar

Como já se viu até agora neste capítulo, a discussão sobre o conhecimento profissional do professor de Matemática não é nova. Durante as últimas décadas surgiram alguns modelos e formas de o apreender. Existe uma abundante investigação acerca do conhecimento do professor de Matemática relativamente a muitos tópicos, mas, no caso do conhecimento do professor que ensina Estatística, a investigação é bastante mais escassa. Isto também acontece pelo facto do tema Estatística no ensino ter um historial muito mais recente, tendo vindo a ser gradualmente introduzido nos diversos níveis dos currículos escolares nacionais e internacionais.

Dado que a Estatística é uma disciplina com uma especificidade própria (Groth, 2007; Moore, 1998; Rossman et al., 2006; Shaughnessy, 2007), vários autores apelam à necessidade de se conhecer caraterísticas intrínsecas do conhecimento do professor necessário para o ensino e aprendizagem nesta área (Burgess, 2006, 2009; Groth, 2007; Shaughnessy, 2007). Burgess (2006) e Groth (2007) assumem que os quadros teóricos do conhecimento profissional do professor de Matemática, por si só, não são apropriados para a Estatística devido às diferenças existentes entre as duas áreas. Em causa está o facto de reconhecerem, tal como outros autores, que a Estatística não é um ramo da Matemática mas uma disciplina independente com métodos e objetivos fundamentais distintos (Batanero et al., 2013; Burrill & Biehler, 2011; Cobb & Moore, 1997; delMas, 2004; Scheaffer, 2006; Shaughnessy, 2006, 2007). Contudo, estes autores consideram que o facto da Matemática ter um papel fundamental na prática estatística pode justificar que a investigação sobre o conhecimento matemático para ensinar possa apoiar, pelo menos como ponto de partida, a investigação sobre o conhecimento do professor no ensino da Estatística. Com base no modelo do *conhecimento matemático para ensinar* de Ball e colaboradores, já referido neste capítulo, Burgess (2006, 2008, 2009) e Groth (2007, 2013) têm procurado criar e desenvolver modelos para analisar e compreender o conhecimento do professor em Estatística, os quais se apresentam de seguida.

3.3.1 O modelo de Groth

O quadro de referência que Groth (2007) apresentou baseia-se, por um lado, no projeto designado por *Aprendendo matemática para ensinar* de Hill e colaboradores (2004; 2005) e, por outro, num modelo para a resolução de problemas estatísticos apresentado no relatório *GAISE* preparado por Franklin et al. (2007), publicado pela Sociedade Americana de Estatística. O modelo do *conhecimento estatístico para ensinar* (Groth, 2007) integra apenas duas componentes do *conhecimento de conteúdo* do modelo do *conhecimento matemático para ensinar* (Hill, Schilling & Ball, 2004; Hill, Rowan & Ball, 2005) – o *conhecimento comum* e *o conhecimento especializado* – e considera, ainda, as quatro componentes do processo de investigação estatística: (1) formulação de questões; (2) recolha de dados; (3) análise de dados; e (4) interpretação dos resultados (Franklin et al., 2007), embora esse processo não figure no esquema do modelo que o autor apresentou (figura 11). Este modelo também não inclui o *PCK*, nem nenhuma das suas componentes.

Trata-se de um modelo que visa o conhecimento estatístico do professor (que consta na figura 11 como conhecimento *não matemático*), e assume também que o conhecimento *matemático* é igualmente mobilizado no ensino da Estatística. A figura 11 abaixo ilustra uma estrutura do *conhecimento estatístico para ensinar* segundo Groth (2007):

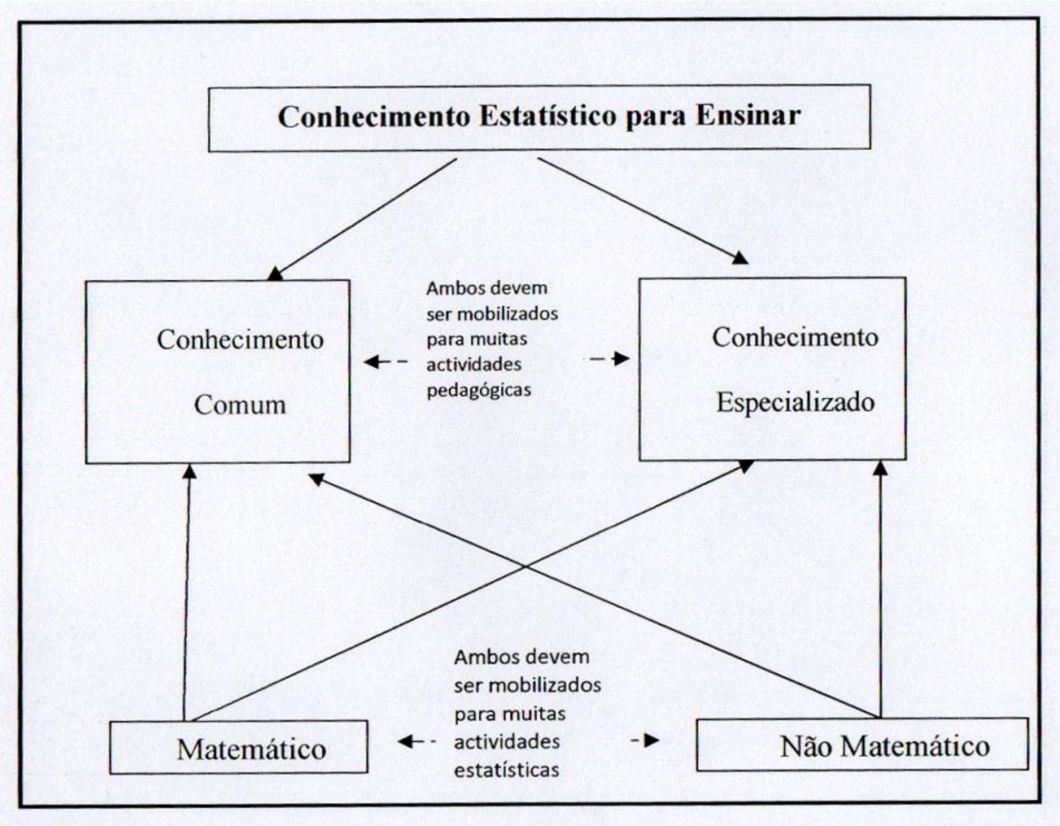


Figura 11: O conhecimento estatístico para ensinar (adaptado de Groth, 2007, p. 429).

Neste modelo, Groth (2007) destaca o *conhecimento comum* e o *conhecimento especializado* do conteúdo, tal como definido por Hill e colaboradores (2004), como componentes do *conhecimento estatístico para ensinar*. Nas palavras do autor, estes conhecimentos são descritos do seguinte modo:

> *Conhecimento comum* diz respeito a competências desenvolvidos em cursos convencionais de Matemática, tais como calcular com precisão, fazer afirmações matemáticas corretas e resolver problemas. *Conhecimento especializado* é desenvolvido tendo cuidadosamente em conta questões e dilemas matemáticos que surgem nos contextos de ensino. Relaciona-se com tarefas tais como fornecer explicações compreensíveis, apreciar os métodos não convencionais de resolução de problemas dos alunos e construir e avaliar múltiplas representações dos conceitos. (Groth, 2007, p. 428)

Para o autor, tanto o *conhecimento comum* como o *conhecimento especializado* podem ser importantes para cada uma das fases do processo de investigação estatística; contudo, dependendo do tipo de tarefa estatística proposta na sala de aula, esta pode exigir que estes conhecimentos (comum ou especializado) sejam de cariz matemático ou estatístico. Adicionalmente, Groth (2007) refere que muitas atividades estatísticas evocam esta combinação de conhecimentos estatísticos e matemáticos e não a sua compartimentação. Por isso, nesta mesma classificação, muitas vezes tem o cuidado de mencionar se uma tarefa irá requerer principalmente conhecimento matemático ou conhecimento não matemático, ou seja, estatístico.

No quadro seguinte (quadro 3) apresenta-se exemplos de algumas situações, que podem surgir na fase de formulação de questões do processo de investigação estatística, e que utilizam principalmente um conhecimento matemático ou estatístico, comum ou especializado, de acordo com Groth (2007).

Quadro 3: Aspetos do conhecimento estatístico para ensinar
(adaptado de Groth, 2007, p. 430)

Formulação de questões	Conhecimento comum	Conhecimento especializado
Conhecimento matemático	Fazer uma leitura correta dos dados apresentados no problema através de diagramas tais como: diagrama de pontos, diagrama de caule e folhas, diagrama de extremos e quartis, caixa de bigodes (boxplots), etc.	Compreender como é que os alunos interpretam os diferentes tipos de gráficos estatísticos.
Conhecimento não matemático (estatístico)	Compreender a distinção entre uma questão determinista e uma questão de natureza inferencial.	Compreender a repercussão das questões formuladas no processo investigativo e avaliar as questões colocadas pelos alunos.

No quadro 4, apresentam-se alguns exemplos de situações que podem surgir na fase de análise de dados do processo de investigação estatística, e que mobilizam principalmente conhecimento matemático ou estatístico, comum ou especializado, de acordo com Groth (2007).

Quadro 4: Aspetos do conhecimento estatístico para ensinar
(adaptado de Groth, 2007, p. 430)

Análise de dados	Conhecimento comum	Conhecimento especializado
Conhecimento matemático	Calcular a média, moda e mediana.	Identificar as propriedades da média, da moda, da mediana, em que os alunos possam ter dificuldades de compreensão.
Conhecimento não matemático (estatístico)	Proceder à escolha da medida de tendência central que melhor representa um conjunto de dados. Determinar a média de um conjunto de dados proveniente de um contexto de *signal in noise*.[11]	Ter a noção de que os alunos que tendem a calcular a média aritmética de um conjunto de dados sem pensar na sua relevância no contexto dos dados.

Apesar de Groth (2007) destacar, no quadro do *conhecimento estatístico para ensinar*, apenas subdomínios do *conhecimento do conteúdo* (Hill, Schilling & Ball, 2004; Hill, Rowan & Ball, 2005), designadamente o conhecimento comum e o conhecimento especializado, parece que o conhecimento especializado, tal como o autor se lhe refere, para além de incluir questões e dilemas matemáticos que possam surgir no contexto escolar, engloba também questões que sugerem uma certa familiaridade com os alunos e os seus raciocínios nas tarefas estatísticas propostas. Por exemplo, destacam-se, das duas tabelas acima, duas situações matemáticas que envolvem um conhecimento especializado: uma que consiste em compreender como é que os alunos leem ou interpretam diferentes tipos de gráficos estatísticos e outra que diz respeito à identificação de propriedades das medidas de tendência central em que os alunos possam ter mais dificuldades de compreensão. Este *conhecimento especializado* de Groth parece também incluir aspetos pedagógicos e relativos a alunos, mais concretamente, parece compreender aspetos do *conhecimento de conteúdo e dos alunos* (Ball et al., 2005, 2008). Ball et al. (2005, 2008) alertam que pode haver situações em que é difícil discernir entre *conhecimento especializado* e *conhecimento de conteúdo e*

[11] *Signal in noise* é uma expressão que pode ser usada na Estatística para designar um ponto de referência dentro da variabilidade dos dados.

dos alunos. Este foi um dos motivos que levou Groth (2013) a expandir o modelo do *conhecimento estatístico para ensinar* à vertente pedagógica e também a assumir a necessidade de se identificar mais aspetos que facilitem a distinção entre *conhecimento especializado* e *conhecimento de conteúdo e alunos*. Para Groth (2007), o desenvolvimento do *conhecimento estatístico para ensinar* passa pela preparação de cursos/programas que ajudem os futuros professores ou professores no ativo a desenvolverem o *conhecimento comum* e o *conhecimento especializado* para o ensino da Estatística, os quais não devem deixar de incluir exemplos ou situações que provenham do contexto escolar para serem exploradas e estudadas por esses professores.

Numa publicação mais recente, Groth (2013) propõe um quadro teórico para caraterizar o *conhecimento estatístico para ensinar*, que refina e amplia o seu trabalho anterior (Groth, 2007). Este novo modelo adiciona à dimensão do *conhecimento de conteúdo*, o *conhecimento do horizonte* (Ball et al., 2008), e passa também a incluir a dimensão do *conhecimento pedagógico de conteúdo*, com as suas três categorias habituais (Ball et al., 2008). Integra, ainda, dois construtos que não existiam no modelo anterior, mas que existem na literatura associadas à Educação Matemática, *compreensões chave do desenvolvimento* e *ideias pedagogicamente poderosas* (Groth, 2013, citando Simon (2006) e Silverman & Thompson (2008), respetivamente).

Neste novo modelo, o construto *compreensões chave do desenvolvimento* é um mecanismo que identifica assuntos ou situações de referência particularmente úteis para a aprendizagem, que irão permitir desenvolver o *conhecimento do conteúdo* do professor; já o construto *ideias pedagogicamente poderosas* é um mecanismo que permite identificar assuntos ou situações relevantes para o desenvolvimento do *conhecimento pedagógico* do professor (Groth, 2013). Estes dois construtos relacionam-se quando o professor usa uma ideia ou conteúdo chave do seu conhecimento (comum, especializado ou do horizonte) e o transforma adequadamente no ensino, tornando-o compreensível para os alunos. Neste modelo, o *conhecimento de conteúdo e dos alunos* é considerado fundamental na formação das *ideias pedagogicamente poderosas* dado que o professor, antes de lecionar uma ideia chave, deve começar por ponderar quais as diferentes estratégias usadas pelos alunos e suas dificuldades (Groth, 2013). A apresentação deste modelo integra alguns exemplos empíricos que foram recolhidos de uma disciplina de licenciatura frequentada por futuros professores do ensino básico (que pretendia desenvolver o *conhecimento estatístico para ensinar* desses alunos); os dados surgiram de respostas escritas produzidas por esses futuros professores a tarefas diversificadas que Groth (2013) propôs no decurso da disciplina, sob a sua responsabilidade.

Ao relacionar os dados empíricos que reuniu com o modelo que desenvolveu, o autor concluiu, por exemplo, que selecionar e relacionar representações gráficas que expõem e condensam os dados como apoio à compreensão de gráficos que ocultam alguns dos seus dados é uma *ideia chave* que o professor ou futuro professor deve possuir na componente do *conhecimento especializado*. Adicionalmente, analisar, interpretar e avaliar representações gráficas não convencionais (nomeadamente, produzidas por alunos) é também um exemplo de uma *ideia chave* que deve estar presente ou ser desenvolvida na componente do *conhecimento especializado* do professor (Groth, 2013).

Com respeito ao *conhecimento comum*, o autor apontou uma *ideia chave* que tem a ver com o professor entender, de um ponto de vista geral, o contributo da teoria das Probabilidades para a Estatística, mais concretamente, como é que elas se podem relacionar quando se lida com dados empíricos. Uma ideia chave do *conhecimento do horizonte* que Groth (2013) indicou foi influenciada pelo trabalho de Garfield e Ben-Zvi (2008), e diz respeito ao professor saber quais as noções básicas que deve desenvolver junto dos alunos do ensino básico antes de eles trabalharem a noção de desvio-padrão de uma forma mais formal. O autor recomenda que as ideias chave que detetou e outras que a investigação venha a revelar devem ser desenvolvidas na formação inicial do futuro professor, quer na componente de conteúdo quer na componente pedagógica.

3.3.2 O modelo de Burgess

Para Burgess (2011), no ensino da Estatística, pretende-se que os alunos alcancem o raciocínio e o pensamento estatístico, indo-se além da literacia estatística básica. Este autor menciona ainda que os professores precisam de adquirir experiência e compreensão do processo investigativo estatístico (Burgess, 2006). Um ponto de partida para se abordar uma investigação estatística na sala de aula é fornecer os dados aos alunos e, quer o professor, quer os alunos, ou ambos decidirem qual a questão/problema a ser solucionada; outra possibilidade é dar liberdade aos alunos para conduzirem uma investigação mais aberta. Este tipo de abordagem permite que os alunos se foquem nas fases de análise de dados e conclusões do processo investigativo. À medida que os alunos vão ganhando mais experiência, podem começar o ciclo investigativo de outra forma, com um problema, questão ou com uma hipótese e, de seguida, passar à recolha de dados, análise, interpretação e conclusões (Burgess, 2006, 2008).

Burgess (2006) propõe o quadro teórico *Conhecimento estatístico para ensinar* "para examinar o conhecimento do professor tal como é usado na prática letiva da estatística" (p. 1). Trata-se de um modelo bidimensional que pretende perscrutar o conhecimento do professor na sala de aula, no ensino da Estatística através de investigações estatísticas. À semelhança do modelo de Groth (2007; 2013), o modelo de Burgess (2006) foi construído com base nos trabalhos de Hill, Schilling e Ball (2004) e Ball, Thames e Phelps (2005) sobre o *conhecimento matemático para ensinar*, tomando em consideração algumas das suas componentes do *conhecimento de conteúdo* e do *conhecimento pedagógico de conteúdo* e tem a particularidade de articular essas componentes com o modelo do pensamento estatístico de Wild e Pfannkuck (1999). Burgess (2006) destaca, do trabalho dos primeiros autores, apenas quatro das seis componentes que integram o modelo do *conhecimento matemático para ensinar*, nomeadamente: o *conhecimento comum* e o *conhecimento especializado* da vertente do conhecimento do conteúdo; o *conhecimento do conteúdo e dos alunos* e o *conhecimento do conteúdo e do ensino* da vertente do *conhecimento pedagógico de conteúdo*. Do trabalho dos últimos autores, Burgess tem em conta o modelo de quatro dimensões do "pensamento estatístico no questionamento empírico" (Wild & Pfannkuck, 1999, p. 1): (i) o ciclo investigativo (que inclui: problema, plano, dados, análise e conclusões); (ii) tipos de pensamento estatístico; (3) o ciclo interrogativo (que inclui: gerar, procurar, interpretar e julgar); e (4) disposições (que incorporam: ceticismo e imaginação).

Este modelo de Wild e Pfannkuck (1999) compreende os aspetos principais do pensamento estatístico (reconhecimento da *necessidade de dados*; *transnumeração*; *consideração da variação*; *raciocínio com modelos*; *integração da Estatística e do contexto*) e aspetos do pensamento mais gerais que podem ser considerados como parte integrante da resolução de problemas (ciclo investigativo, ciclo interrogativo e disposições). Os aspetos principais do pensamento estatístico são descritos do seguinte modo (Burgess, 2006; Shaughnessy, 2007; Wild & Pfannkuch, 1999):

- *Reconhecimento da necessidade de dados*: em vez de se basear em indícios anedóticos e tomada de consciência de quanto mais dados melhores conclusões se poderão tirar;
- *Transnumeração*: capacidade de mudar de representação para chegar a uma melhor compreensão dos dados;
- *Consideração da variação*: capacidade de medir, modelar e controlar a variação e decisões sobre os dados; envolve também procurar e descrever padrões de variação e compreendê-los em relação ao contexto;
- *Raciocínio com modelos*: partir dos modelos mais simples para os mais complexos;
- *Integração da Estatística e contexto*: esta conexão é uma componente fundamental do pensamento estatístico, envolvendo a capacidade de relacionar o modelo estatístico usado com o contexto do problema. Em situações mais elementares envolve, por exemplo, interpretar medidas estatísticas e representações gráficas à luz do seu contexto (Shaughnessy, 2007).

Na literatura em Educação Estatística estes cinco aspetos do pensamento estatístico também surgem muitas vezes identificados como aspetos do raciocínio estatístico, uma vez que são também aspetos relevantes e necessários para assegurar o desenvolvimento do raciocínio estatístico (e.g., Batanero et al., 2013; Godino et al., 2008).

O esquema seguinte (quadro 5) ilustra o quadro teórico de Burgess (2006, 2008) para examinar o conhecimento do professor mobilizado na sala de aula, no ensino da Estatística, através de investigações. Com este modelo o autor pretende descrever o conhecimento necessário para ensinar Estatística na sala de aula e, além disso, compreender que aspetos do conhecimento deste quadro teórico são mais ou menos relevantes para a prática do professor.

Quadro 5: Modelo de Burgess (adaptado de Burgess, 2006, p. 4)

			Conhecimento estatístico para ensinar			
			Conhecimento do conteúdo		**Conhecimento pedagógico do conteúdo**	
			Conhec. Comum do conteúdo	**Conhec. especializado do conteúdo**	**Conhec. do conteúdo e dos alunos**	**Conhec. do conteúdo e do ensino**
Pensamento Estatístico no Questionamento Empírico	**Tipos de Pensamento**	**Necessidade de dados**				
		Transnumeração				
		Variação				
		Raciocínio com modelos				
		Integração da Estatística e contexto				
	Ciclo investigative					
	Ciclo interrogative					
	Disposições					

Burgess (2008) usou o modelo do conhecimento estatístico para ensinar na observação de dois professores do ensino básico com pouca experiência profissional enquanto ensinavam Estatística através de investigações. A análise da prática destes dois professores mostrou que a maioria dos aspetos do *conhecimento* (do *conteúdo* e *pedagógico do conteúdo*) e do *pensamento estatístico* é fundamental na prática. Num estudo de continuidade do anterior, Burgess (2009) realiza uma investigação com quatro professores, em que o autor, além usar a observação de aulas, recorre também a entrevistas, para aceder às explicações das opções tomadas pelos professores nas aulas e à sua reflexão sobre os episódios de aula.

Neste trabalho, o autor reafirma a utilidade do quadro teórico usado na descrição do conhecimento do professor em relação ao pensamento estatístico. Revela também alguns exemplos de oportunidades perdidas pelos professores participantes nas suas aulas devido à falta de *conhecimento pedagógico*. O professor com mais dificuldade neste domínio foi o que pareceu ter um *conhecimento do conteúdo* mais fragilizado, por exemplo, quanto ao uso de representações e capacidade de integrar *Estatística e* *contexto* com argumentos válidos. A análise de situações em que o professor desperdiçou a oportunidade de um ensino ou de uma aprendizagem mais produtiva revelou também que uma parte essencial do trabalho do professor, no contexto das investigações estatísticas, diz respeito a dar resposta às questões dos alunos, avaliar as suas respostas e também orientá-los nesses trabalhos. Para o autor o desenvolvimento profissional na área da Estatística deverá ter em consideração as dificuldades detetadas na investigação, para que a aprendizagem não fique comprometida (Burgess, 2008, 2009).

Oliveira e Henriques (2014) usaram o quadro teórico de Burgess (2009) para analisar o *conhecimento estatístico para ensinar* de futuros professores do ensino básico e secundário. Neste estudo as autoras não contemplaram algumas dimensões do modelo de Burgess (2009), tais como o *conhecimento comum* e as *disposições*. O *conhecimento estatístico para ensinar* foi analisado a partir de reflexões escritas de dez futuros professores, que frequentaram uma disciplina de didática da Matemática lecionada pelas autoras, que incluía um módulo dedicado ao ensino e aprendizagem da Estatística e Probabilidades. Essas reflexões foram realizadas individualmente ou a pares e incidiram em episódios de aula sobre uma investigação estatística que envolvia dados de censos da população de três países.

Neste estudo, os futuros professores revelaram capacidade em avaliar as produções de alunos em termos de *conhecimento estatístico* envolvido, identificando elementos da maioria dos aspetos do pensamento estatístico. A única componente que não emergiu nessas análises foi a da *variação*. O facto do conceito de *variação* ser novo para os futuros professores e das características dos episódios selecionados poderem limitar a mobilização desse conhecimento, visto incorporar populações em vez de amostras, são razões apontadas, pelas autoras, como possíveis justificações para a ausência de dados que articulassem *conhecimento especializado de conteúdo* e a *variação* nas respostas dos futuros professores. Relativamente ao *conhecimento pedagógico*, os dados revelam que a dimensão do *conhecimento do conteúdo e do ensino* tem uma maior expressão do que a do *conhecimento do conteúdo e dos alunos*, com os futuros professores a evidenciarem uma maior capacidade para interpretar a atuação do professor atendendo aos aspetos do pensamento estatístico. Tal como aconteceu com a *variação* no domínio do *conhecimento do conteúdo*, este aspeto do pensamento estatístico também não surgiu nas reflexões dos futuros professores no domínio do *conhecimento pedagógico*, o que pode estar relacionado com as razões apontadas para o caso do *conhecimento de conteúdo* e com a falta de experiência de lecionação. Adicionalmente, os resultados deste estudo revelam uma maior capacidade dos futuros professores em assinalar aspetos relativos à *transnumeração* e *raciocínio com modelos* na análise da intervenção dos alunos nos episódios de aula (Oliveira & Henriques, 2014).

Segundo Oliveira e Henriques (2014), o quadro concetual é adequado para analisar o *conhecimento estatístico para ensinar* do futuro professor em contexto de formação inicial, embora reconhecendo algumas dificuldades em dissociar algumas componentes do conhecimento do professor, nomeadamente, entre *conhecimento especializado* e *conhecimento de conteúdo e alunos,* tal como a literatura indica. As autoras consideram que através deste modelo pode-se analisar o *conhecimento estatístico para ensinar* de futuros professores, identificando-se aspetos a melhorar, e pode-se ainda reunir situações a explorar no âmbito da didática da Estatística.

3.4 O Conhecimento didático do professor em Estatística

O estudo desenvolvido no presente trabalho foca-se essencialmente na prática letiva do professor, adotando-se como modelo o *conhecimento didático* de Ponte e Oliveira (2002) e Ponte (2012), que incide no estudo do conhecimento do professor em ação. O conhecimento do professor de Matemática é considerado dinâmico em vez de fixo ou estático (Burgess, 2006; Chapman & Ponte, 2006; Petrou & Goulding, 2011; Ponte, 2012; Rowland et al., 2009), daí que a investigação sobre a prática é relevante para aceder-se à natureza do conhecimento do professor e ao modo como ele se desenvolve (Beswick et al., 2012; Burgess 2006; 2009; Depaepe et al., 2015; Ponte & Oliveira, 2002; Ponte 2012; Rowland et al., 2008). Depaepe et al. (2013) salientam que estudos baseados numa perspetiva situada do conhecimento do professor de Matemática podem dar um contributo significativo para o campo da Educação Matemática e ter a função de complementar, nalguns aspetos, os resultados de outros estudos realizados nesta área tais como os centrados numa perspetiva cognitiva, em que frequentemente analisam o fenómeno em causa através de testes, sem ter em conta o contexto de trabalho e outros fatores que intervêm diretamente ou indiretamente no trabalho do professor.

Tendo em mente que o desenvolvimento do conhecimento do professor é um processo lento, Chapman (2013) considera fundamental que a investigação na área do conhecimento do professor de Matemática vá além daquilo que este consegue ou não fazer, compreendendo de forma profunda como é que os professores dão sentido a este conhecimento. Assim sendo, o pensamento do professor, bem como aspetos afetivos do professor, tais como intenções e motivações, devem também ser contemplados na investigação (Chapman, 2013; Depaepe et al., 2015). Adicionalmente, Beswick et al. (2012) defendem que é importante, quando se analisa o conhecimento do professor através de um conjunto de categorias ou de uma forma compartimentada, complementar-se essa análise com uma visão holística deste conhecimento, o que permitirá capturar aspetos que, de outro modo, poderiam permanecer ocultos, uma vez que o fenómeno em causa é complexo.

Speer, King e Howell (2015) consideram que há necessidade de se examinar a concetualização do *conhecimento matemático para ensinar* quando usado na investigação do conhecimento profissional do professor de níveis de ensino que não sejam o básico. Estes autores enfatizam que os professores do ensino básico e do ensino secundário têm *backgrounds* diferentes, nomeadamente ao nível dos conteúdos matemáticos, e que o trabalho matemático requerido aos professores desses níveis de ensino é também muito diferente. Por isso, alertam que uma reconcetualização do *conhecimento comum* e do *conhecimento especializado do conteúdo* do modelo do *conhecimento matemático para ensinar* (Ball et al., 2008) seria uma mais-valia para a coerência e robustez da investigação ao nível do ensino secundário e superior. Mesmo ao nível do ensino básico, na Estatística, Burgess (2009) considera que essa reconcetualização é necessária visto que o *conhecimento comum* facilmente é associado a conhecimentos requeridos ao nível da literacia estatística básica. Porém, a literatura atual mostra precisamente várias lacunas no conhecimento do professor nesse domínio.

Adicionalmente, na Educação Estatística, há autores que atribuem ao *conhecimento comum* e ao *especializado* as mesmas situações ou capacidades. Por exemplo, Groth (2007) diz que proceder à escolha da medida de tendência central ou de localização que melhor representa um conjunto de dados é *conhecimento comum*, e essa mesma situação é *conhecimento especializado* para Burgess (2009). Também Speer et al. (2015) argumentam que o *conhecimento especializado*, tal como é descrito habitualmente, é um conhecimento que não é só usado por professores de Matemática mas também por matemáticos que o usam rotineiramente enquanto investigadores. Dados esses factos, parece que a distinção entre os conhecimentos de conteúdo, nomeadamente, do *conhecimento comum* e *especializado* tal como Ball et al. (2008) os definem (como um conhecimento básico, não específico do professor; como um conhecimento específico do professor, respetivamente) necessita de reformulação. No contexto da Estatística, por exemplo, há necessidade de um certo consenso no que se entende por literacia estatística e avaliar se é apropriado fazê-la corresponder ao *conhecimento comum*, e de considerar quanto ao *conhecimento especializado* que este deve atender não só ao conhecimento de Estatística do professor mas também às exigências de cada nível de ensino. Também Groth (2013) indica a necessidade de uma maior clarificação do ponto de vista concetual entre *conhecimento especializado* e *conhecimento do conteúdo e dos alunos*.

Ao nível do ensino secundário, as orientações curriculares americanas e a investigação na Educação Estatística visam o desenvolvimento do raciocínio estatístico dos alunos e enfatizam a necessidade de mudança de paradigma no ensino. Assim sendo, essa literatura sugere um ensino que explore dados reais, que contemple análises de dados que considere a variabilidade dos dados e que integre a realização de interpretações em torno das representações gráficas e dos resultados à luz dos dados e do contexto de cada situação (e.g., Burgess, 2009; Burrill & Biehler, 2011; Curcio & Artzt, 1996; Garfield & Ben-Zvi, 2008; Shaughnessy et al., 2009). Também é sugerido que se pondere a compreensão conceptual dos conceitos, suas propriedades e relações entre conceitos, bem como a integração da tecnologia (e.g., Burrill & Biehler, 2011; Franklin et al., 2007; Garfield & Ben-Zvi, 2005, 2008; Shaughnessy, 2006; 2007). Isto constitui um verdadeiro desafio para o professor cuja experiência estatística se baseia sobretudo na estatística descritiva (Henriques & Ponte, 2014; Pfannkuch & Ben-Zvi, 2011) e em ensiná-la como um conjunto de cálculos e procedimentos, em que os dados e o seu contexto e variabilidade não são devidamente enfatizados (e.g., Ben-Zvi & Garfield, 2004; Burrill, 2008; delMas & Liu, 2005; Engel & Sedlmeier, 2011).

Vários estudos empíricos no campo do ensino e aprendizagem da Estatística referidos nesta tese focam-se no conhecimento do aluno (e.g., Boaventura & Fernandes, 2004; Cobb et al., 2003; Sorto et al., 2011; Watson & Shaughnessy, 2004) ou nalguma área específica do conhecimento profissional do professor relativamente a um determinado assunto estatístico (e.g., Groth & Bergner, 2006; Ribeiro et al., 2009; Watson & Callingham, 2013). Começam também a surgir alguns estudos que pretendem não só cobrir mais aspetos do domínio do conhecimento profissional do professor em Estatística mas também considerar tópicos que abranjam vários assuntos estatísticos tais como o processo investigativo e a variabilidade (e.g., Burgess, 2009, 2010; Groth, 2007; Oliveira & Henriques, 2014). Há um longo caminho a percorrer no domínio do conhecimento profissional do professor e seu desenvolvimento na Educação Estatística, também por ser uma área relativamente recente e com poucas investigações centradas no ensino secundário (Burrill, 2008; Fernandes et al., 2011; Shaughnessy, 2007).

Neste estudo, ao usar-se o modelo do *conhecimento didático* não houve necessidade de realizar-se a distinção entre *conhecimento comum* e *conhecimento especializado*; contudo na dimensão de *conhecimento de Estatística* do *conhecimento didático* procurou-se identificar e compreender, de um modo geral, conhecimentos considerados como pertinentes para o ensino da Estatística e também os conhecimentos que os professores revelam acerca das ideias estatísticas fundamentais (Batanero et al., 2013; Burrill, 2008; Burrill & Biehler, 2011). O modelo adotado é inspirado no modelo do *conhecimento didático do professor* de Matemática de Canavarro (2003), Ponte e Oliveira (2002) e de Ponte (2012) que, como foi referido, é também influenciado por Shulman (1986), mas em que se assume que este se trata de um conhecimento orientado para as situações da prática e que considera o seu desenvolvimento.

Este modelo do *conhecimento didático*, que é constituído por quatro domínios intimamente relacionados com a prática letiva, é descrito tendo em conta a especificidade da Estatística e do seu ensino, tal como apontado por vários autores da literatura nessa área. Dá atenção a aspetos do raciocínio estatístico (Batanero et al., 2013; Burgess, 2009; Shaughnessy, 2006; Wild & Pfannkuch, 1999) que se interligam com os conceitos estatísticos fundamentais no contexto do nosso país (nomeadamente, dados e amostragem, variabilidade, associação e modelação) que devem ser desenvolvidos pelos alunos no ensino secundário (e.g., Batanero et al., 2013; Burrill & Biehler, 2011; Garfield & Ben-Zvi, 2008). Apresentam-se em seguida as quatro dimensões consideradas e que, embora apresentadas isoladamente, se encontram fortemente interrelacionadas (ver figura 12).

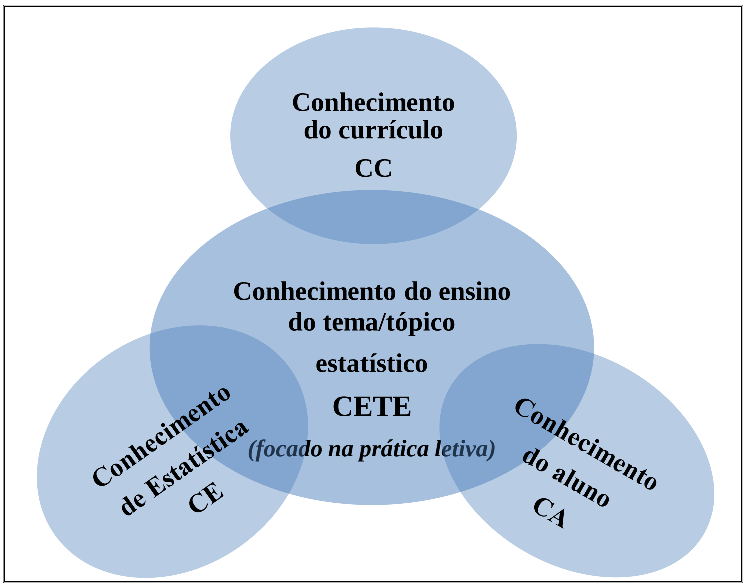


Figura 12: O modelo do conhecimento didático adotado no estudo.

O *conhecimento de Estatística* (CE) abarca também conhecimento de Matemática que o professor possui e usa para levar os alunos a compreender o que fazem e por que o fazem. Inclui conceitos, procedimentos, representações, explicações e perspetiva geral sobre a Estatística e as suas conexões internas e externas. Diz respeito ao conhecimento que o professor tem dos aspetos específicos do saber que ensina (Godino et al., 2008; Shaughnessy, 2007), no caso tendo em conta a sua singularidade relativamente à Matemática. Ou seja, envolve, em particular, o conhecimento de conceitos estatísticos fundamentais, suas interligações e conhecimento dos diversos dilemas e desafios que envolvem lidar com essas ideias (Batanero et al., 2013; Burrill & Biehler, 2011).

O *conhecimento do currículo* (CC) diz respeito ao conhecimento dos objetivos e finalidades do currículo escolar e sua articulação vertical e horizontal, incluindo ainda o conhecimento de diferentes recursos que podem ser usados para ensinar Estatística e de possíveis formas de avaliação. Este conhecimento tem um papel fundamental na tomada de decisões sobre os assuntos a que se deve dedicar mais tempo ou aprofundar e sobre a forma de orientar o processo de ensino-aprendizagem da Estatística (Godino et al., 2008; Ponte & Oliveira, 2002; Ponte, 2012), na disciplina de Matemática.

O *conhecimento do aluno* (CA) diz respeito a uma compreensão sobre os interesses dos alunos, necessidades e dificuldades de aprendizagem frequentes, bem como o conhecimento de aspetos de cariz social que possam influenciar o desempenho dos alunos (Godino et al., 2008; Ponte & Oliveira, 2002; Ponte, 2012). Inclui a capacidade de prever erros e obstáculos que os alunos enfrentam na aprendizagem (Godino et al., 2008; Groth, 2013). Integra, ainda, o conhecimento de estratégias, raciocínios, erros e dificuldades na aprendizagem, quando os alunos lidam com conceitos e com as tarefas propostas (Groth, 2013; Henrique & Ponte, 2014; Ponte & Oliveira, 2002; Ponte, 2012) e a capacidade de compreender e avaliar as respostas e o trabalho dos seus alunos (Burgess, 2008; Groth, 2013).

O *conhecimento do ensino do tema/tópico estatístico* (CETE) diz fundamentalmente respeito à condução efetiva das situações de aprendizagem (Ponte & Oliveira, 2002; Ponte, 2012), procurando-se compreender o modo como cada tópico estatístico foi desenvolvido e com que grau de exigência (Godino et al., 2008). O conhecimento do professor revela-se nas suas opções quanto às tarefas a propor aos alunos, as interações que ocorrem em aula e a regulação geral do processo de ensino-aprendizagem, mas atende-se também à própria atividade de planificação, que antecede, assim como às reflexões do professor sobre a sua prática (Godino et al., 2008; Henriques & Ponte, 2014; Ponte & Oliveira, 2002; Ponte, 2012). Esta dimensão do conhecimento abarca também a sua capacidade de ajudar os alunos a desenvolver o raciocínio e pensamento estatísticos (Scheaffer, 2006; Shaughnessy, 2007). Esta ocupa o núcleo fundamental do *conhecimento didático* visto ser essencialmente focada na prática, dada a sua centralidade e interações com as outras dimensões do modelo (Ponte, 2012).

Tal como está definido, este modelo pode contribuir para a compreensão do conhecimento didático do professor em Estatística, no ensino secundário, de uma forma abrangente e holística, tal como vários autores recomendam (e.g., Beswick et al., 2012; Chapman, 2013; Depaepe et al., 2015). Também, a investigação em Educação que considera o conhecimento do professor como dinâmico e procura capturar mais informação sobre este conhecimento poderá contribuir para pensar e refletir sobre ações com vista ao desenvolvimento da formação inicial e da formação contínua que combine teoria e prática (e.g., Beswick et al., 2012; Burgess, 2009; Chapman, 2013; Groth, 2013; Henriques & Oliveira, 2013; Oliveira & Henriques, 2014; Ponte, 2011, 2014; Rowland, 2013; Rowland et. al, 2008).

CAPÍTULO 4

METODOLOGIA

Neste capítulo são apresentadas e justificadas as opções metodológicas adotadas à luz da problemática e das questões que norteiam o estudo. Inclui-se, nomeadamente, a escolha da abordagem de investigação, os critérios de seleção dos participantes, as questões éticas consideradas, os métodos de recolha de dados, bem como o modo como os dados empíricos recolhidos foram analisados.

4.1 A natureza do estudo

O presente estudo centra-se no professor, com especial incidência na sua prática profissional, visando caraterizar o seu conhecimento didático em Estatística, na disciplina de Matemática A, quando leciona o tema no 10.º ano do ensino secundário. Tendo em conta a perspetiva assumida neste trabalho relativamente ao conhecimento didático do professor, concebido como um conhecimento fortemente ancorado na prática letiva e condicionado pelos contextos em que este desenvolve a sua prática, afigura-se essencial aceder também aos significados atribuídos pelos professores às suas ações no seu contexto natural de prática profissional. Assim sendo, a opção por realizar um estudo de natureza qualitativa de cunho interpretativo parece particularmente adequada. Neste paradigma, o trabalho do investigador baseia-se na variabilidade das relações comportamento e significado, produzidos pelos indivíduos nas suas interações sociais, focando-se no modo como se desenvolvem e mantêm estes sistemas de significado e não simplesmente nos comportamentos observáveis (Léssard-Hebert, Goyette & Boutin, 1994). Deste modo, o objetivo do investigador é analisar e compreender a ação dos indivíduos, a qual engloba os comportamentos e os significados que o indivíduo lhes atribui e, ainda, os significados conferidos por aqueles que interagem com este indivíduo (Erickson, 1986).

De acordo com Bogdan e Biklen (1994), em Educação, a investigação qualitativa pode assumir muitas formas e ser realizada em múltiplos contextos. Denzin e Lincoln (1998) definem a investigação qualitativa como sendo um multimétodo que envolve uma perspetiva interpretativa e naturalista face ao seu objeto de estudo. Os investigadores qualitativos estudam a realidade no seu contexto natural, procurando dar-lhe sentido, interpretando os fenómenos de acordo com os significados que os participantes do estudo lhes atribuem (Savenye & Robinson, 2005). Bogdan e Biklen (1994) enunciam cinco caraterísticas fundamentais da investigação qualitativa, apesar de considerarem que algumas dessas caraterísticas possam estar mais ou menos presentes, ou até inexistirem, num estudo dessa natureza: (1) a fonte direta dos dados é o ambiente natural, constituindo o investigador o principal instrumento; (2) a investigação é descritiva; (3) para o investigador, os processos merecem uma maior ênfase do que os resultados ou produtos; (4) o foco da investigação está nos significados atribuídos pelos indivíduos; e (5) a análise dos dados é feita de forma indutiva.

No presente estudo, uma das principais fontes dos dados é a sala de aula, o contexto natural onde decorre a prática profissional das professoras participantes, sendo a investigadora o principal instrumento da recolha de dados na forma de observadora participante; a escola é também o lugar ou o contexto escolhido para a realização de entrevistas às professoras. Com efeito, de acordo com Bogdan e Biklen (1994), um investigador qualitativo tende a recorrer o mais possível ao contexto onde decorre a ação dos participantes do estudo, pois este é determinante, pela influência que exerce sobre o comportamento humano.

Neste trabalho pretende-se dar maior ênfase aos processos do que aos resultados, dado que o foco do estudo está na compreensão e descrição dos comportamentos das professoras a partir da sua perspetiva no seu ambiente natural, não tendo a investigadora a intenção de exercer qualquer controlo sobre o contexto natural ou sobre os comportamentos das participantes do estudo, tal como refere Patton (2002). Deste modo, os dados recolhidos e a sua análise assumem uma natureza descritiva.

Além disso, a análise de dados é feita de forma indutiva, não tendo como objetivo confirmar ou provar hipóteses, mas antes, gerar hipóteses e produzir conhecimentos específicos para o estudo em causa. Indo ao encontro do que referem Bogdan e Biklen (1994) como caraterística da investigação qualitativa, a análise de dados vai sendo realizada à medida que o investigador desenvolve o processo de recolha de dados e os examina, e com o refinamento contínuo das questões e opções metodológicas do estudo.

4.2 Os estudos de caso

Esta investigação de natureza qualitativa segue a modalidade de estudo de caso. Para Ponte (2006), os estudos de caso têm sido usados na investigação em Educação Matemática para indagar uma variedade de questões, nomeadamente, do conhecimento e das práticas profissionais dos professores. Vários autores consideram que, mais do que uma escolha metodológica, o estudo de caso é um *design* de investigação (Merriam, 1988; Ponte, 2006; Stake, 1994, 2007). Este *design* é para Merriam (1988) adequado para a compreensão e interpretação dos fenómenos que ocorrem no contexto educacional. Adicionalmente, Yin (2003) carateriza o estudo de caso como “um estudo empírico que investiga um fenómeno contemporâneo, num contexto real, especialmente quando a fronteira entre o fenómeno e o contexto não está claramente definida” (p. 13). Mais especificamente, para Merriam (1988) e Yin (1989), a realização de estudos de caso é particularmente relevante quando: (i) se pretende responder a questões de natureza explicativa do tipo “como?” e “porquê?”; (ii) não existe qualquer intenção de se controlar os acontecimentos; e (iii) se pretende descrever, analisar e explicitar de modo holístico um fenómeno contemporâneo em estudo, como produto final.

Merriam (1988) acrescenta, ainda, como aspeto a ser ponderado na adoção do estudo de caso como modalidade de investigação, a identificação de um sistema limitado, ou seja, de um caso sobre o qual a investigação vai incidir. De acordo com esta autora, um estudo de caso qualitativo tem as seguintes propriedades essenciais: é particularístico, descritivo, heurístico e indutivo. Assume um caráter particularístico, dado debruçar-se sobre uma situação específica (o caso); é descritivo, pois pretende realizar uma descrição detalhada do fenómeno em estudo; é heurístico, uma vez que se foca na compreensão do fenómeno em estudo; é indutivo pelo facto de se apoiar no raciocínio indutivo. As generalizações, conceitos ou hipóteses emergem da análise dos dados, a que se tem acesso direto, podendo proporcionar a descoberta de novos significados.

O principal objetivo do presente estudo é investigar o conhecimento didático do professor, compreendendo aspetos deste conhecimento, fundamentalmente a partir da análise da sua prática, explicitando-o também de um modo holístico. Assim, a opção por realizar estudos de caso é justificada por se pretender estudar um fenómeno contemporâneo no seu contexto natural, pelas caraterísticas das questões de investigação e pela ausência de controlo da investigadora sobre os acontecimentos.

Stake (1994, 2007) apresenta três tipos de estudos de caso que têm por base diferentes interesses, por parte do investigador: intrínsecos, instrumentais e coletivos. O investigador que realiza um estudo de caso intrínseco pretende focar-se num acontecimento em particular. O investigador tem, sobretudo, um interesse em compreender aspetos específicos do caso, e não a partir dele compreender aspetos sobre outros casos ou uma situação mais geral. O investigador adota o estudo de caso instrumental quando visa alcançar mais do que a compreensão daquele caso particular, contribuindo para aprofundar o conhecimento de um determinado fenómeno. O estudo de caso coletivo refere-se a um estudo instrumental que se vai estendendo a vários casos, semelhantes ou distintos. Estes casos são escolhidos com base no entendimento de que cada um deles poderá dar um contributo importante para uma melhor compreensão do fenómeno em estudo.

No que diz respeito ao propósito principal dos estudos de caso em Educação, Merriam (1988) refere que estes podem ser fundamentalmente: descritivos, analíticos e avaliativos. Os estudos de caso descritivos visam descrever o que se pretende estudar, sem que haja uma preocupação analítica; nesta categorização estão incluídos os chamados estudos de caso exploratórios (Ponte, 2006) que servem para fornecer uma informação preliminar do caso. Os estudos de caso analíticos pretendem teorizar sobre o fenómeno em estudo. A partir da análise dos dados do estudo e sua interpretação, é possível problematizar o seu objeto e dar um contributo adicional para a teoria. Os estudos de caso avaliativos visam, para além de descrever e explicar o caso, emitir um juízo avaliativo. Ponte (2006) considera que, de entre os estudos de caso do tipo exploratórios, descritivos e analíticos, "são os estudos de cunho analítico que proporcionam um mais significativo avanço do conhecimento" (p. 6).

Na presente investigação opta-se pela realização de dois estudos de caso instrumentais (Stake, 2007) de modo a incluir-se mais evidência e diversidade de informação, pretendendo-se ir além da compreensão de cada caso. Pretende-se também realizar estudos de caso comparativos, de forma a ser possível salientar diferenças ou semelhanças entre os dois casos. Cada um destes estudos de caso instrumentais é também analítico, pois pretende-se não só descrever o mais fielmente possível o conhecimento didático do professor em Estatística, mas também realizar uma análise profunda que contribua para o desenvolvimento teórico desta temática, ou seja, o conhecimento didático do professor que ensina o tema Estatística na disciplina de Matemática, no ensino secundário.

4.3 Participantes no estudo

4.3.1 O contexto de seleção dos participantes

A identificação e escolha dos casos são passos importantes na realização de estudos de caso (Merriam, 1988; Yin, 1989). No presente estudo, cada caso diz respeito a uma professora de Matemática do ensino secundário. Através de cada um deles não há o intuito de generalizar resultados, mas entender o fenómeno em estudo de uma forma aprofundada, procurando-se ir além de cada caso constituído (Stake, 2007).

Uma opção assumida na escolha dos participantes do estudo foi a de procurar professores de Matemática do ensino secundário em escolas secundárias da região do grande Porto (zona de residência da investigadora) que se disponibilizassem para participar na investigação. Dado que nem sempre é fácil que os professores se disponham a abrir a sua sala de aula a uma pessoa estranha e a despender do seu tempo para realizar encontros com a investigadora, uma opção que se considerou mais profícua foi desenvolver uma ação de formação, no âmbito da Estatística, num centro de formação que abrangesse várias escolas secundárias do Grande Porto e, a partir desta, selecionar os professores participantes. Este contexto da ação de formação iria, por um lado, permitir ter um contacto mais próximo com professores interessados na sua formação profissional na área da Estatística e, por outro, facilitar o acesso à sua prática letiva.

A ação de formação realizada, na modalidade de oficina de formação, da responsabilidade da investigadora, foi designada por “Abordagens e tarefas para o ensino da Estatística no ensino secundário”. Os seus objetivos eram os seguintes: (i) atualizar e aprofundar conhecimentos na área da Estatística lecionada no ensino secundário; (ii) atualizar conhecimentos na área da didática da Matemática, com particular incidência na didática da Estatística; (iii) incentivar a utilização da calculadora gráfica, da folha de cálculo (Excel) e de vários *sites* fidedignos da internet para a exploração de tarefas em sala de aula, no âmbito do tema da Estatística no ensino secundário; (iv) desenvolver a capacidade de elaboração de tarefas para a aula de Estatística que proporcionem aos alunos o desenvolvimento do seu raciocínio estatístico; e (v) promover a reflexão partilhada incidindo especialmente na implementação em sala de aula de tarefas produzidas pelos formandos com vista ao desenvolvimento do raciocínio estatístico dos alunos.

Esta ação decorreu num centro de formação da região Norte, de janeiro a abril de 2011, com nove sessões presenciais que corresponderam a 25 horas, e outras 25 horas não presenciais (ver quadro 6). Participaram 19 professores de Matemática do ensino básico e secundário, de oito escolas com ensino secundário do Grande Porto em que apenas quatro deles não tinham qualquer experiência de ensino da Estatística no ensino secundário.

Nas duas sessões iniciais da formação foram discutidos com os formandos alguns conceitos e potencialidades das investigações estatísticas na promoção do raciocínio estatístico dos alunos, com exemplos retirados da literatura em Educação Estatística (Franklin et al., 2007; Scheaffer, 2006; Shaughnessy & Chance, 2005). Alguns fundamentos teóricos de acordo com o programa do ensino secundário vigente na altura (ME, 2001) foram também considerados nessas sessões e aprofundados ao longo da formação. Recorreu-se também à utilização da tecnologia, como a folha de cálculo e a calculadora gráfica, para que os formandos desenvolvessem algumas competências necessárias para o seu uso e alguma familiaridade com a utilização dos menus de Estatística das ferramentas tecnológicas, nomeadamente, com um leque de representações numéricas e gráficas que poderiam apoiar as análises de dados que iriam realizar nas sessões posteriores.

Nas cinco sessões seguintes da formação foram apresentadas tarefas de cariz investigativo que incidiam nas fases de análise de dados (reais) e de interpretação e comunicação de resultados do processo de investigação estatística (Shaughnessy et al., 2009). Cada uma dessas cinco sessões organizou-se em duas partes principais. Na primeira parte, os formandos trabalharam em grupos de dois ou três elementos na tarefa proposta, fundamentalmente, nas duas fases de uma investigação estatística referidas; essa parte era concluída depois de se fazer uma súmula das noções e representações mais úteis usadas pelos formandos na análise de dados, bem como uma síntese das conclusões obtidas. Na segunda parte, eram analisados diálogos e interações na sala de aula entre professores e alunos na exploração das mesmas tarefas abordadas pelos formandos, sob a forma de episódios escritos, retirados de Shaughnessy et al. (2009) e de Shaughnessy e Chance (2005), com o objetivo principal de discutir aspetos didáticos no âmbito do ensino da Estatística. Adicionalmente, era realizada uma súmula das representações usadas e conclusões obtidas nos excertos analisados, que era ainda confrontada com a súmula feita após a exploração dos dados das tarefas originais pelos formandos, de modo a refletir-se sobre o potencial da tarefa no aprofundamento dos conhecimentos estatísticos dos alunos.

Na penúltima sessão os formandos fizeram uma apresentação final do trabalho de grupo[12] que lhes foi proposto, que incluiu uma discussão final sobre os trabalhos realizados. Este trabalho consistiu na criação e implementação de uma tarefa de cariz investigativo com dados de contextos reais, e que fossem significativos para os alunos, numa turma de um dos formandos do grupo. Nessa apresentação, os grupos indicaram a tarefa de cariz investigativo que propuseram a uma turma (de 8.º, 9.º ou 10.º ano) e fizeram uma síntese da análise de dados e conclusões efetuadas pela turma e ainda uma reflexão global relativa à produção e implementação da tarefa (desafios) e também comentários à análise e resultados obtidos. Com base nessas apresentações finais foram

[12] Para tornar esse trabalho exequível, optou-se por formar grupos de professores que lecionavam nas mesmas escolas, tendo sido formados três grupos de três elementos e cinco grupos de dois elementos.

elaborados oito posters que foram apresentados num *workshop* que foi divulgado pelo Centro de Formação a professores de Matemática dos ensinos básico e secundário, de todas as escolas abrangidas por este. Este *workshop* ocorreu nas instalações do referido centro e constituiu a última sessão da ação de formação.

Quadro 6: Síntese cronológica da ação de formação

Data (número de sessões)	Síntese de atividades
Janeiro (uma sessão)	Apresentação. Alguns fundamentos teóricos básicos. Estatística escolar com a folha de cálculo e calculadora gráfica.
Fevereiro (duas sessões)	Estatística escolar com a folha de cálculo e calculadora gráfica. 1ª tarefa - *Dados demográficos de três países.*
Março (três sessões)	2ª tarefa - *Erupções do geiser americano Old Faithful.* 3.ª tarefa - *Atletismo olímpico.* 4.ª tarefa - *Preços de casas numa localidade*.
Abril (três sessões)	5.ª tarefa - *Clientes do Starbucks, um estudo de observação*. Apresentação dos trabalhos dos formandos. *Workshop* final.

Como foi referido acima, a ação de formação foi sobretudo o contexto privilegiado para se escolherem os participantes do estudo, não havendo a intenção de se avaliar o impacto desta ação na prática dos professores que constituíram os estudos de caso. Contudo, ao longo da recolha de dados, a investigadora esteve sempre atenta a aspetos da ação de formação que as professoras relevassem, assim como recursos ou ideias que tenham manifestado ter intenção de levar para as suas aulas e a forma como os concretizaram e aplicaram nas aulas, caso o tenham feito. Uma vez que a compreensão do conhecimento didático do professor que ensina Estatística, a partir da sua prática, corresponde ao objetivo principal do presente estudo, é importante ter em conta todas as experiências potencialmente significativas para as professoras, neste âmbito.

4.3.2 As participantes escolhidas

Ao se propor a ação de formação referida, esperava-se que todos os professores lecionassem pelo menos uma turma de Matemática A, no 10.º ano, no momento em que decorreu a formação; contudo, a maioria dos professores que se inscreveu lecionava, nesse ano letivo, apenas turmas do 3.º ciclo do ensino básico. De facto, deste grupo de 19 professores apenas três professoras lecionavam a disciplina de Matemática A no 10.º ano.

Os critérios definidos, originalmente, para a seleção dos professores participantes foram os seguintes: (i) terem experiência no ensino secundário, dado que se pretendia assegurar que as suas práticas e perspetivas no âmbito da Estatística, deste nível de escolaridade, fossem mais significativas e robustas; (ii) estarem a lecionar nesse ano letivo, pelo menos uma turma de 10.º ano de Matemática A; (iii) pertencerem à mesma escola ou escolas com uma localização próxima e lecionarem em horários que permitissem que a observação de aulas pudesse ser feita pela investigadora, o maior número de vezes possível, a uma turma de cada professora.

As três professoras que satisfaziam todos os requisitos assumidos no estudo eram: Carmo, Estela e Lia (pseudónimos). Carmo e Lia lecionavam numa mesma escola secundária e Estela noutra não muito distante. Após a conclusão da ação de formação, as três professoras foram contactadas pela investigadora para averiguar da sua disponibilidade para colaborar na investigação.

Assim que lhes foi explicado o objetivo do estudo, o processo de recolha de dados com os tempos requeridos e como é que esses dados iriam ser utilizados e divulgados, as três professoras indicaram o seu interesse e o seu consentimento para participarem na investigação. Adicionalmente, cada uma delas também considerou que as turmas de 10.º ano que lecionavam reuniam condições favoráveis para a realização da investigação. Assim que Carmo deu a sua anuência, também informou que não garantia que conseguisse evitar faltar a algumas aulas em que tinha previsto lecionar o tema da

Estatística, por motivos de saúde. Assim, à partida, decidiu-se recolher dados com as três professoras e apenas após terminada essa recolha tomar uma decisão quanto ao número de casos a considerar no estudo, uma vez que se a professora Carmo não pudesse lecionar todas as aulas isso poderia comprometer a exequibilidade do estudo por ausência de dados. A professora Carmo acabou por faltar apenas a uma aula.

No final da recolha de dados realizou-se uma análise preliminar de dados para se atender à diversidade e contraste dos casos. Dessa análise, há a referir que Estela destacou-se das outras duas professoras por revelar, por um lado, possuir mais experiência no ensino do tema da Estatística e, por outro, dar maior ênfase no ensino à utilização da calculadora gráfica; adicionalmente, Estela e Lia revelaram, relativamente a Carmo, um maior à vontade em refletir sobre a Estatística escolar, sobretudo nas entrevistas. Também, o facto de se ter reunido uma quantidade considerável de dados para cada professora, levantou a questão de que se optasse por considerar os três casos e todos esses dados fossem contemplados na investigação a tese ficaria muito extensa. Por conseguinte, esta análise preliminar dos dados obtidos contribuiu para que se optasse pela seleção de apenas dois estudos de caso, das professoras Estela e Lia, com a inclusão de todos os dados reunidos para cada uma delas.

As duas professoras selecionadas possuem uma licenciatura em Matemática (ramo educacional) mas Estela realizou também um mestrado em Educação Matemática. No momento de início do estudo, Lia era professora há cerca de 15 anos mas apenas tinha trabalhado com o ensino secundário nas duas últimas escolas que lecionou. Estela possuía mais de 23 anos de experiência, lecionando frequentemente no ensino secundário.

4.4.3 As turmas selecionadas

As professoras selecionadas para o estudo tinham horários completos e lecionavam quer turmas do ensino básico quer do ensino secundário. Relativamente à Matemática A de 10.º ano, ambas lecionavam apenas uma turma. Ambas as turmas eram formadas por 25 alunos. Em termos do género, no caso de Estela a turma incluía 13 raparigas e 12 rapazes e no caso de Lia, a turma incluía 15 raparigas e 10 rapazes. Relativamente ao aproveitamento das turmas, Estela considerava que a sua turma tinha um bom aproveitamento e era, de uma maneira geral, uma das melhores de 10.º ano na escola. Lia, por sua vez, achava que na sua turma apenas se destacavam quatro bons alunos, com resultados relativamente consistentes ao longo do ano; os restantes tinham desempenhos variados entre razoável e fraco. Estela caraterizava a sua turma como trabalhadora e responsável. Lia considerava a sua turma, de uma maneira geral, respeitadora das normas mas pouco participativa nas atividades na aula.

4.4 Questões éticas

Para Bogdan e Biklen (1994), há alguns princípios éticos que acompanham o trabalho da maioria dos investigadores qualitativos. Por exemplo, na negociação do consentimento para se efetuar um estudo de natureza qualitativa, o investigador deve ser claro e explícito relativamente a cada um dos termos do acordo, tratando todos os intervenientes respeitosamente; deve ainda proteger a identidade dos participantes, bem como tratá-los com respeito desde o início até à conclusão do estudo.

Como foi mencionado na secção anterior, os participantes da investigação foram escolhidos entre os formandos que participaram na ação de formação, cuja formadora foi a investigadora, mas apenas após o término da ação de formação é que a investigadora fez o convite às professoras para participarem no estudo, de modo a que não se sentissem coagidas a aceitar participar pelo facto de a investigadora ser a formadora na ação e, consequentemente, avaliadora das mesmas. Há ainda a destacar que esta formação, além de servir como contexto importante de acesso ao campo, identificando potenciais participantes para o estudo investigativo, também constituiu uma das formas de a investigadora dar uma contrapartida aos potenciais participantes do estudo e, de modo mais alargado, aos outros professores não participantes, reconhecendo a responsabilidade social que acompanha a realização de uma investigação em Educação (Bogdan & Biklen, 1994).

Houve também o cuidado de explicar, às professoras participantes, o objetivo do estudo de natureza qualitativa que se planeava realizar, e o que lhes seria requerido em termos de colaboração (Bogdan & Biklen, 1994). Nomeadamente, foram informadas sobre onde e quando ocorreria a recolha de dados, as opções planeadas quanto aos métodos de recolha de dados, o modo como os dados seriam registados (áudio-gravação para as entrevistas longas; vídeo-gravação das aulas) e divulgados, dando-se a garantia de se preservar a sua identidade e privacidade, bem como dos alunos das suas turmas e das instituições de ensino a que pertenciam.

As professoras entenderam o caráter voluntário da sua participação e a possibilidade de surgirem alterações, ao longo da investigação, sobre os termos do acordo pretendido, sobre as quais seriam consultadas. As professoras contactadas deram consentimento oral para que se usassem tecnologias para os registos de dados e também deram autorização à divulgação dos dados no âmbito da investigação. Depois deste consentimento das professoras, obteve-se autorização adicional dos diretores das escolas, onde as professoras lecionavam, aos quais foi garantido o anonimato de todas as fontes e das suas respetivas instituições na divulgação do trabalho investigativo, através de um documento escrito assinado pela investigadora (anexo 6). A confidencialidade e privacidade dos alunos das turmas das professoras participantes do estudo foram também assumidas junto dos encarregados de educação. Todos eles assinaram uma declaração a autorizar a vídeo-gravação das aulas do seu educando e o uso dos dados recolhidos na aula, no âmbito da investigação (anexo 6). Os alunos das turmas de cada professora também deram o seu consentimento oral para que alguns dos seus registos escritos realizados nos seus cadernos diários ou as representações gráficas que tivessem produzido nas suas calculadoras gráficas fossem filmados no decurso das aulas observadas. Constituiu também um compromisso da investigadora comunicar aos participantes do estudo os resultados da investigação quando esta estivesse concluída.

Este trabalho investigativo seguiu, assim, os objetivos, princípios e orientações de foro ético que devem ser respeitados e fomentados no âmbito da investigação em Educação e Formação, de acordo com a Carta Ética para a Investigação em Educação e Formação[13] do Instituto de Educação da Universidade de Lisboa.

[13] CEIEF - D.R., 2.ª série, N.º 52, 15 de Março de 2016.

4.5 Métodos de recolha de dados

Tendo em conta a problemática do estudo e as opções metodológicas tomadas, neste trabalho investigativo, recorre-se à observação participante que é um método particularmente adequado quando o investigador tem como objetivo "compreender um meio social que à partida lhe é estranho ou exterior, e que lhe vai permitir integrar-se progressivamente nas atividades das pessoas que nele vive" (Lessard-Hébert et al., 1994, p.155). A observação participante foi desenvolvida a partir de estudos etnográficos em antropologia e sociologia e é uma abordagem amplamente reconhecida na investigação qualitativa (Savenye & Robinson, 2005). Também Ludke e André (1986) referem que a observação participante possibilita um contacto mais próximo e pessoal entre o investigador e o fenómeno em estudo, facilitando a compreensão dos factos ocorridos. Vários autores consideram a observação participante como uma abordagem geral que admite a observação direta e momentos de observação sistemática e, ainda, a utilização de uma diversidade de métodos de recolha de dados que pressupõem um envolvimento controlado do investigador junto dos sujeitos e das suas ações, tendo o investigador sempre presente na sua mente o estudo a que se propôs realizar (e.g., Bogdan & Biklen, 1994; Ludke & André, 1986; Savenye & Robinson, 2005).

A recolha de dados neste estudo decorreu entre finais de abril e início de julho de 2011, sendo as seguintes as opções metodológicas instrumentais planeadas, relativamente a cada professora que constituiu um caso: observação de aulas, entrevista, notas de campo e recolha documental. Na investigação qualitativa a recolha de dados é frequentemente realizada com recurso a estes métodos (Merriam, 1988; Savenye & Robinson, 2005).

4.5.1 Observação de aulas

Marshall e Rossman (2006) definem a observação participante como um método de pesquisa qualitativa que oferece oportunidade ao investigador de ver, ouvir e experienciar a realidade próximo do modo como os participantes o fazem. No contexto deste estudo a observação da prática tem um papel destacado. Pretendeu-se que essa observação tivesse um grau de participação limitado, havendo um reduzido envolvimento da investigadora no contexto, por um lado, para que não perturbasse o ambiente natural da sala das professoras e, por outro, no que se refere aos pontos de vista dos professores participantes no estudo, para que não afetasse as ações e decisões das professoras (Bogdan & Biklen, 1994; Evertson & Green, 1986, referidos em Lessard-Hebert et al., 1994). Neste caso, pode falar-se da investigadora como uma "participante periférica" (Adler & Adler, 1994, p. 379), dado que observou e interagiu com os membros do cenário em estudo, de modo a estabelecer uma "identidade de dentro", contudo não participando nas atividades centrais que ocorreram neste cenário.

De acordo com a ordem dos temas matemáticos a lecionar no 10.º ano (Matemática A) sugerida no programa escolar (ME, 2001), o tema da Estatística é habitualmente lecionado no 3.º período. Foi precisamente nesse período que as professoras do estudo lecionaram este tema, pelo que a observação de aulas decorreu durante aproximadamente quatro semanas, com o objetivo geral de recolher dados sobre o conhecimento didático do professor que ensina Estatística, a partir das suas práticas.

A investigadora procurou observar o maior número de aulas possível do tema da Estatística, numa turma de Matemática A de 10.º ano, de cada participante do estudo. A proximidade das duas escolas e compatibilidade dos horários das turmas (Estela lecionava geralmente da parte da manhã e Lia da parte da tarde) agilizou esse processo. Apenas não foram observadas as aulas dedicadas à realização do teste de avaliação que incidiu sobre a Estatística e uma aula de revisões para o teste final do 3.º período porque as professoras informaram que essas revisões não iriam incidir, especificamente, sobre o tema da Estatística. Assim, foram observadas um total de oito aulas para cada professora, entre a última semana do mês de maio e meados do mês de junho de 2011. Os tópicos abordados em cada uma das aulas observadas de cada professora encontram-se nos quadros 7 e 8.

As aulas observadas foram registadas em áudio e vídeo, sendo também efetuadas algumas notas de campo relativamente a comentários efetuados pelas professoras nos momentos antes e pós-aulas observadas, realizados ainda no contexto escolar, geralmente, na sala dos professores de cada escola. Nessas notas de campo também foram assinaladas algumas ideias, interpretações e preocupações relativamente a situações das aulas observadas que chamaram particular atenção à investigadora.

Para além do consentimento para a observação e registo das aulas por parte das professoras, da direção das escolas e dos encarregados de educação dos alunos, já mencionado (anexo 6), a investigadora também teve autorização de cada professora para circular pela sala de aula, bem como dos alunos para filmar alguns dos seus registos escritos. Habitualmente, a investigadora filmou a partir do fundo da sala de aula e procurou sempre que possível aproximar-se da professora quando esta ia ao encontro dos alunos nos seus lugares. Para perturbar o menos possível o ambiente da aula, a investigadora usou calçado com sola de borracha e tentou ter uma postura e aparência discretas.

Quadro 7: Aulas observadas e respetivos tópicos relativamente à professora Estela

Estela	**Aulas Observadas (Obs) – Data**			
Tópicos principais	**Obs 1 - 27/5** População e amostra. Organização e dados. Resolução de tarefas, em grupo, com o apoio da calculadora gráfica.	**Obs 2 - 2/6** Organização e interpretação de dados. Representações gráficas. Resolução de tarefas, em grupo, com o apoio da calculadora gráfica. Função cumulativa.	**Obs 3 - 3/6** Medidas de localização. Quartis. Desvio-padrão. Resolução de tarefas.	**Obs 4 - 7/6** Medidas de localização. *Escola virtual* – dados agrupados e representações.
	Obs 5 - 9/6 Medidas de dispersão. Propriedades da média e do desvio-padrão. Desvio-padrão em gráficos de barras.	**Obs 6 - 13/6** Correlação e regressão linear. Coeficiente de correlação. Resolução de tarefas.	**Obs 7 - 16/6** Regressão linear/ modelo linear. Resolução de tarefas.	**Obs 8 - 17/6** Modelo de regressão linear. Resolução de tarefas.

Quadro 8: Aulas observadas e respetivos tópicos relativamente à professora Lia

Lia	Aulas Observadas (Obs) – Data			
Tópicos principais	**Obs 1 - 23/5** População e amostra. Resolução de tarefas.	**Obs 2 - 24/5** Organização e interpretação de dados: Representações, tabela de frequências e função cumulativa. Resolução de tarefas.	**Obs 3 - 30/5** Organização e interpretação de dados. Medidas de localização. Resolução de tarefas.	**Obs 4 - 31/5** Medidas de localização. Leitura do diagrama de extremos e quartis. Resolução de tarefas.
	Obs 5 - 7/6 Resolução de tarefas.	**Obs 6 - 13/6** Medidas de dispersão. Resolução de tarefas.	**Obs 7 - 14/6** Desvio-padrão. *Escola virtual* - Correlação e regressar linear. Resolução de tarefas.	**Obs 8 - 17/6** Correlação e regressar linear. Resolução de tarefas.

4.5.2 Entrevistas

Segundo Bogdan e Biklen (1994), na investigação qualitativa, a entrevista é muitas vezes utilizada em conjunto com outros métodos de recolha de dados, nomeadamente, com a observação e a recolha documental. Os autores referem que os dados obtidos da entrevista são descritivos "e na linguagem do próprio sujeito, permitindo ao investigador desenvolver intuitivamente uma ideia sobre a maneira como os sujeitos interpretam aspetos do mundo" (p. 134). De acordo com Seidman (2006), na raiz das entrevistas está a intenção de compreender a experiência do professor e o significado que este atribui a essa experiência. Adicionalmente, este autor refere que a entrevista é um instrumento relevante para se obter conhecimento sobre a Educação e sobre outras áreas sociais.

Nesta pesquisa, pretendeu-se recorrer às entrevistas para se aceder aos significados, intenções ou pensamentos das professoras relativamente às suas experiências profissionais e, nomeadamente, face a um conjunto de situações com vista a compreender de uma forma aprofundada os seus conhecimentos didáticos em Estatística. As entrevistas realizadas foram do tipo semi-estruturado, ou seja, basearam-se num guião que incluía um conjunto de tópicos ou questões gerais (de natureza relativamente aberta) já previamente pensados pela investigadora e tiveram uma duração entre 90 minutos a 120 minutos. Estas entrevistas semi-estruturadas, embora tenham sido feitas com base num guião previamente elaborado, tinham a particularidade de permitir uma certa flexibilidade na ordem das questões ou dos temas a abordar e na introdução de novas questões motivadas pelas intervenções do entrevistado (Bogdan & Biklen, 1994), de forma a manter-se uma fluência no diálogo, mas sem se perder de vista a recolha de informação pretendida. Numa entrevista deste tipo se houver alguma informação que não tenha ficado suficientemente esclarecida há sempre a possibilidade de a retomar numa entrevista posterior dado que numa investigação que integra a entrevista qualitativa a informação é cumulativa (Bogdan & Biklen, 1994).

No presente estudo realizaram-se três entrevistas semi-estruturadas com cada participante. Estas entrevistas incidiram sobre diferentes aspetos e foram áudio-gravadas e, após a sua realização, transcritas na íntegra pela investigadora. A primeira entrevista foi realizada antes das professoras iniciarem a lecionação do tema da Estatística, a segunda foi realizada um pouco antes do término das aulas e a terceira imediatamente após o término das aulas. Estas três entrevistas visavam contribuir com elementos relevantes para responder às questões de investigação.

Na primeira entrevista (E1), além de procurar-se conhecer aspetos mais gerais do conhecimento didático das professoras em Estatística, pretendia-se também conhecer alguns aspetos relativos ao conhecimento curricular, as opções para a unidade de Estatística que iam iniciar e caraterísticas das tarefas que propunham habitualmente nas suas aulas (anexo 1). Na segunda entrevista (E2) procurou-se recolher dados acerca do conhecimento didático de cada professora em Estatística, solicitando-lhes, por exemplo, algumas reflexões em torno de algum conteúdo ou da sua atuação ou ainda acerca da atuação de algum aluno relativamente a uma seleção abrangente e também variada de situações e materiais oriundos das práticas de cada professora nos diferentes tópicos que abordaram (anexo 1). À semelhança da segunda entrevista, a terceira (E3) teve também como propósito continuar a aprofundar a temática do conhecimento didático do professor em Estatística, nos aspetos específicos contemplados nas questões do estudo e também para fazer um balanço final da unidade de Estatística lecionada (anexo 1).

As entrevistas decorreram essencialmente com base nas questões dos guiões preparados e incluíram alguns materiais diversificados como estratégia para melhor compreender as ideias das professoras à luz das questões do estudo (quadro 9). Os guiões elaborados para cada uma das professoras tinham, de uma maneira geral, uma estrutura semelhante; no entanto, também por assinalarem assuntos específicos das práticas de cada uma das professoras, deram origem a algumas questões diferenciadas para cada professora. Em todas as entrevistas as professoras tinham ao dispor o respetivo manual adotado, o programa de Matemática A (versão integral e versão resumida num quadro elaborado pela investigadora), uma calculadora gráfica e papel e lápis, caso sentissem necessidade de escrever para concretizarem alguma ideia.

A marcação das entrevistas foi feita de acordo com as disponibilidades das professoras. A investigadora procurou que as entrevistas das duas professoras fossem marcadas em dias diferentes, para poder refletir sobre as entrevistas realizadas e ponderar sobre eventuais aspetos a melhorar (quadro 10).

Quadro 9: Materiais usados nas entrevistas longas

E1	Tarefas (anexo 2), manuais adotados.
E2	Situações (anexo 2), materiais apresentados nas aulas, tarefas propostas nas aulas, produções dos alunos.
E3	Tarefas propostas nas aulas e excertos de gravações de aulas.

Quadro 10: Calendarização das entrevistas longas

	E1	**E2**	**E3**
Estela	5 de maio	3 de junho	30 de junho
Lia	6 de maio	7 de junho	1 de julho

Como já foi referido acima, um dos propósitos da entrevista inicial (E1) era caraterizar, globalmente, as tarefas que as professoras utilizam frequentemente nas aulas de Estatística de acordo com as suas experiências profissionais. Por conseguinte, optou-se nessa entrevista inicial, por apresentar às participantes três tarefas diversificadas (designadas por *Média sobre média*, *Tempo de espera pelo autocarro*, *Como está a evoluir a nossa população?* - incluídas no anexo 2), como estratégia para garantir, por um lado, alguns dados que pudessem ser comparados entre elas, independentemente dos exemplos que apresentassem serem díspares ou não; e, por outro lado, para que as professoras fossem levadas a expressarem-se sobre tarefas que potencialmente promovem o raciocínio estatístico dos alunos. Assim, pediu-se a cada professora para indicar se as três tarefas selecionadas lhes eram familiares e também para ponderarem a hipótese de as proporem nas suas aulas.

A tarefa *Média sobre média* proposta na entrevista foi retirada do manual escolar de Costa e Rodrigues (2010, p. 178). Este era o manual adotado na escola de Lia mas não na escola de Estela; contudo, era-lhe bastante familiar, pois tinha por hábito consultar manuais destes autores. Esta tarefa foi selecionada pela investigadora sobretudo por envolver o cálculo de uma média ponderada e por poder causar alguma dificuldade aos alunos (Boaventura & Fernandes, 2004; Callingham, 1997). Este tipo de tarefa, de acordo com Scheaffer (2006), pode ser categorizado como uma tarefa de foco matemático.

A tarefa *Tempo de espera pelo autocarro* foi também retirada do manual escolar de Costa e Rodrigues (2010, p. 232) mas teve uma pequena adaptação na última alínea em que lhe foi acrescentado um pedido de interpretação do resultado do desvio-padrão obtido. Este tipo de tarefa, que envolve, em particular, leitura de um gráfico e algum cálculo, é bastante comum nos manuais escolares, embora o pedido adicional feito não o seja. É também um tipo de tarefa particularmente útil para desenvolver a capacidade de compreensão de gráficos (Curcio & Artzt, 1996) e alguma análise da variabilidade dos dados (Garfield & Ben-Zvi, 2008), tendo potencial para promover o raciocínio estatístico (Franklin et al., 2007).

A tarefa *Como está a evoluir a nossa população?* foi construída pela investigadora e inclui dados reais, que correspondem à distribuição das idades da população portuguesa (agrupadas em classes etárias), em 1991 e 2001. É uma tarefa de foco estatístico que envolve dados e um contexto real e em que a questão formulada só pode ser cabalmente respondida após uma análise de dados adequada (Scheaffer, 2006). Esta tarefa centra-se essencialmente nas fases de análise de dados e de interpretação e comunicação dos resultados do processo de investigação estatístico (Burgess, 2006, 2011). Inclui, adicionalmente, um conjunto de orientações para o trabalho escrito (nomeadamente, procura de padrões/relações nos dados; análise da variabilidade dos dados, reflexão acerca do contexto dos dados e conjeturas/inferências acerca dos dados; comparação das distribuições; uso de representações numéricas e gráficas na análise dos dados e interpretação dos resultados à luz da situação problemática).

Existem outras duas situações, que estão também referenciadas no anexo 2, que incidem em respostas hipotéticas de alunos, criadas pela investigadora, que foram apresentadas, na entrevista 2, às professoras do estudo para que comentassem se as aceitariam como válidas, apresentando as suas razões.

Imediatamente após cada uma das três entrevistas, a investigadora elaborou notas de campo sintéticas sobre o que experienciou no decurso de cada entrevista, no contexto escolar. Essa informação poderia ser útil para refletir sobre os dados do estudo qualitativo, mais especificamente, sobre o modo com estes podiam afetar o plano de investigação e o modo como a investigadora interagia com eles (Bogdan & Biklen, 1994), constituindo um complemento à recolha de dados.

Para além das três entrevistas longas, no período de observação de aulas, antes e após cada aula, foram realizadas entrevistas de curta duração (até 15 minutos), com cada professora, com registo através de notas de campo. Através destas curtas entrevistas, a investigadora pretendia conhecer o objetivo global de cada aula e compreender as opções assumidas e interpretações das professoras face aos acontecimentos mais marcantes ocorridos.

4.5.3 Recolha documental

A recolha documental é um método relevante na investigação uma vez que permite obter informações que podem complementar ou enriquecer os dados obtidos por outras vias. Yin (1989) refere a importância de recolher informação a partir da análise de documentos que possam ser disponibilizados. Existem documentos que, apesar de não terem sido concebidos para o estudo, podem constituir uma fonte de recolha de dados que permitem confrontar as interpretações realizadas com dados provenientes de outras fontes. Foram recolhidos todos os materiais produzidos ou utilizados pelas professoras participantes para as aulas em que lecionaram o tema da Estatística, designadamente, planificações do tema da Estatística, fichas de trabalho, testes sumativos e outros materiais distribuídos aos alunos (em particular, *powerpoints*). Procedeu-se também à recolha das produções escritas dos alunos na resolução de algumas tarefas selecionadas.

4.6 Análise de dados

Seguindo Huberman e Miles (1994), procurou-se entrelaçar a recolha de dados com a sua análise, de modo a que a análise dos primeiros dados recolhidos, interagindo com os objetivos da investigação e os contributos da literatura (teórica e empírica), pudesse ajudar a melhorar as recolhas seguintes, com vista a uma compreensão cada vez mais aprofundada da situação em estudo. A análise de dados é "o processo de busca e de organização sistemática de transcrições de entrevistas, de notas de campo, e de todos os materiais que foram sendo acumulados" (Bogdan & Biklen, 1994, p. 205), com vista a facilitar o entendimento desses materiais e de se encontrar formas de transmitir esse conhecimento.

Huberman e Miles (1994) indicam um modelo interativo (ver figura 13) para a análise de dados composto por três subprocessos: (i) redução de dados – que consiste na seleção, delimitação, simplificação, abstração e transformação do material compilado; (ii) apresentação dos dados – que diz respeito ao tratamento dos dados, condensando-os e representando-os; e (iii) interpretação e verificação das conclusões – que incide na atribuição de significado aos dados reduzidos, expondo-se relações nos dados, explicações e configurações possíveis. De acordo com estes autores, estes subprocessos estão ligados uns aos outros e ocorrem antes da recolha de dados, durante a fase de tomada de decisão do *design* do estudo e sua planificação, durante a recolha de dados (com análises iniciais de alguns dados), bem como depois da recolha de dados para se alcançar os resultados finais.

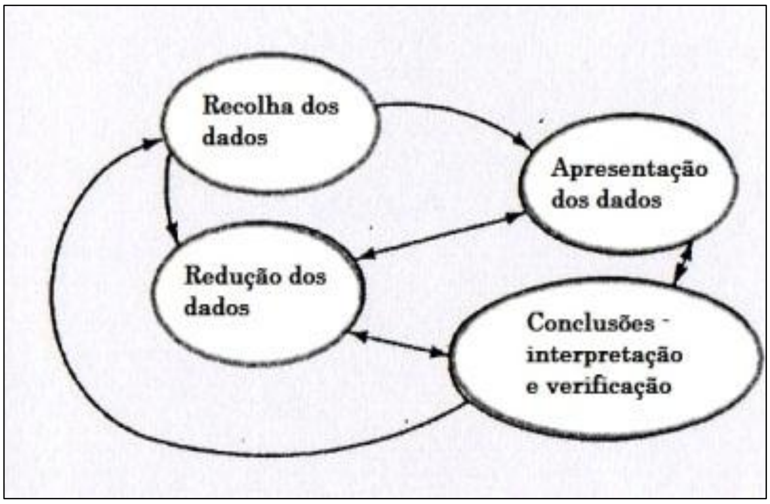


Figura 13: Esquema do modelo interativo de análise de dados (adaptado de Huberman & Miles, 1994, p. 428).

No estudo, os dados foram analisados de forma descritiva e interpretativa, através de um sistema de categorias, que não estava totalmente definido à *priori*. O estudo assume algumas dimensões de análise que foram identificadas com base no objetivo e nas questões do estudo e dentro destas, subdimensões resultantes da revisão de literatura teórica e empírica e de outras que emergiram durante a análise dos dados (ver quadros 11 e 13).

A apresentação dos casos foi feita de forma descritiva expondo um conjunto de dados, onde consta alguma interpretação da investigadora. O processo de comparação dos dois casos iniciou-se após a análise do primeiro caso e assim que os diferentes subcapítulos do segundo caso foram finalizados, o que também permitiu que se fossem aperfeiçoando ambos os casos. Após a descrição e análise de cada caso (capítulo 5 e capítulo 6) é que se concluiu a comparação destes casos e realizou uma discussão mais alargada dos resultados, a qual está incluída nas conclusões do estudo (capítulo 7).

Quadro 11: Dimensões e subdimensões de análise pré-estabelecidas para cada caso

PERCURSO E CONTEXTO PROFISSIONAL	**PERSPETIVA GERAL**	**A ESTATÍSTICA NA SALA DE AULA**
- Escolha da profissão - A escola e relação com colegas e alunos - Participação em projetos e formações - A experiência pessoal na Estatística - A turma e o contexto letivo	- Programa da disciplina de Matemática A - Planificação da unidade de Estatística -Caraterísticas das tarefas estatísticas	- População e amostra - Medidas de localização - Medidas de dispersão - Dados bivariados e regressão linear - Organização e interpretação de dados - Raciocínio estatístico (aspeto transversal que inclui: Dados e contexto, *Transnumeração*, Variabilidade, Raciocínio com modelos, Integração da Estatística e contexto)

Cada caso das professoras foi organizado num capítulo de acordo com as três dimensões de análise (*Percurso e contexto profissional*, *Perspetiva geral sobre o conhecimento didático do professor em Estatística* e *A Estatística na sala de aula*) e suas subdimensões. Houve o cuidado de garantir que a estrutura dos casos ficasse o mais semelhante possível para facilitar a análise comparativa entre eles, o que não impediu que se atendesse a aspetos diferentes ou contrastantes do conhecimento didático de cada professora no âmbito da Estatística. Os títulos dos três subcapítulos principais de cada caso são assim baseados nessas dimensões assumidas, ou seja, o primeiro é *Percurso e contexto profissional*, o segundo é *Perspetiva geral sobre o conhecimento didático da professora em Estatística* e o terceiro é *A Estatística na sala de aula*.

Há a destacar que o último subcapítulo principal (*A Estatística na sala de aula*) centra-se fundamentalmente na prática do professor que ensina o tema Estatística, acabando por ser um dos maiores dos três por incidir na análise de dados provenientes de todos os métodos usados no estudo (ver os código atribuídos, no quadro 12, às entrevistas e observações de aulas). Adicionalmente, este subcapítulo foi subdividido pelos tópicos do programa escolar que as professoras abordaram nas aulas, tendo o cuidado de manter-se a sequenciação dos tópicos usada por cada professora, que coincide com a sugerida pelo programa escolar.

Quadro 12: Códigos inseridos na recolha de dados

Códigos usados nos métodos de recolha de dados	
Entrevista longa	E1, E2 e E3
Observação de aulas	Obs1, Obs2, Obs3, Obs4, Obs5, Obs6, Obs7 e Obs8

Na análise e interpretação dos dados de cada tópico estatístico contemplado deu-se enfase às tarefas e situações[14] abordadas nas aulas (anexos 3 e 4), bem como às respetivas interações (sobretudo entre professora e alunos) proporcionadas relativamente a essas situações/tarefas, uma vez que as práticas das duas professoras se desenvolveram essencialmente em torno da resolução das tarefas propostas. Assim, as tarefas e situações apresentadas neste trabalho foram escolhidas por evidenciarem aspetos centrais do conhecimento didático relativamente ao tópico estatístico tratado na aula.

Todo o trabalho analítico e interpretativo realizado em cada caso e também na discussão do conhecimento didático das duas professoras e na apresentação das conclusões foi orientado pelas questões do estudo, pelas dimensões e subdimensões da análise estabelecidas (quadros 11 e 13) que diziam respeito aos elementos mais marcantes da literatura teórica e empírica sobre os assuntos temáticos desta investigação e também por subdimensões (quadro 13) que emergiram dos dados empíricos recolhidos. Assim, a estruturação e organização de cada caso e da própria discussão mais alargada apresentada nas conclusões foi conduzida por esses mesmos elementos. Adicionalmente, na elaboração das conclusões também se procurou responder às questões de investigação com interpretações o mais abrangentes possíveis.

[14] A maioria destas situações corresponde a uma seleção de representações gráficas e tabulares que as professoras integraram na aula nalguns momentos expositivos.

Quadro 13: Subdimensões adicionais estabelecidas

Subdimensões estabelecidas na discussão alargada do conhecimento didático em Estatística das professoras
Conhecimento do currículo (CC) *Leitura do programa escolar* **Conhecimento do ensino do tópico estatístico (CETE)** *Preparação do tópico* - Sequenciação dos conteúdos - Escolha de tarefas *Condução do ensino* -Compreensão dos conceitos e dos dados -Estabelecimento de conexões envolvendo conceitos e representações **Conhecimento do aluno (CA)** *Expetativas em relação à aprendizagem* *Entendimento acerca do conhecimento desenvolvido pelos alunos* **Conhecimento de Estatística (CE)** *Compreensão de conceitos e dos dados* *Relações estabelecidas envolvendo conceitos e representações*

CAPÍTULO 5

O CASO DE ESTELA

Este capítulo é dedicado ao caso da professora de Matemática Estela. É um capítulo constituído por três partes principais, *Percurso e contexto profissional*, *Perspetiva geral sobre o conhecimento didático do professor em Estatística* e *A Estatística na sala de aula*. Na primeira é realizada uma breve apresentação da professora, descrito o seu percurso e contexto profissional, destacando-se aspetos marcantes da escolha da sua profissão e da relação que possui com colegas e alunos no contexto escolar. É também dada atenção à sua participação em projetos e formações e à sua experiência na Estatística enquanto aluna e formanda. Adicionalmente, apresenta-se uma breve caraterização da turma onde foi acompanhado o trabalho da professora.

Na segunda parte apresenta-se uma perspetiva geral do *conhecimento didático em Estatística* da professora relativamente a alguns aspetos, nomeadamente, à forma como o tema Estatística do programa escolar do ensino secundário é perspetivado pela professora e à leitura que a professora faz dos conteúdos a abordar à medida que percorre os tópicos principais. Nesta segunda parte é, ainda, caraterizado o modo como Estela habitualmente procede para planificar a unidade de Estatística, destacando os materiais e recursos em que se apoia nessa preparação, e tenta-se compreender o que entende por tarefa estatística. Esta parte é rematada com uma discussão guiada pelos domínios do modelo do conhecimento didático em Estatística.

Na terceira parte analisa-se o conhecimento didático da professora em Estatística, tendo por base a sequência de tópicos estatísticos que abordou com a sua turma de 10.º ano do secundário, nomeadamente *População e amostra*; *Organização e interpretação de dados*; *Medidas de localização, Medidas de dispersão* e *Dados bivariados e regressão linear*. Esta parte, além de incluir uma discussão de cada tópico guiada pelos domínios do modelo do conhecimento didático, também incorpora uma discussão, com a mesma organização da realizada por tópicos, mas focada na unidade de Estatística, numa perspetiva geral. Ainda assim, há a referir que os dados dessas discussões apresentam-se mais sistematizados, relacionados e aprofundados no capítulo das conclusões do estudo (capítulo 7), o qual tem a particularidade de integrar uma discussão alargada do conhecimento didático em Estatística das duas professoras participantes do estudo.

5.1 Percurso e contexto profissional

5.1.1 Breve apresentação da professora

Estela é professora de Matemática do 3.º ciclo e do ensino secundário, é casada e tem três filhos. Ingressou na faculdade, numa universidade pública do Norte na sua primeira opção no início da década de 80, no curso de Matemática do ramo educacional, curso este que incluía um estágio integrado com duração de um ano letivo. É uma professora que leciona há mais de 23 anos. Durante a sua carreira docente lecionou em três escolas diferentes pertencentes à DREN, tendo tido a oportunidade de ensinar todos os níveis, desde o 7.º até ao 12.º ano. Concluiu recentemente o mestrado na área da Educação Matemática, numa universidade pública do Centro.

5.1.2 A escolha da profissão

Estela considera que o seu interesse por querer ser professora talvez se deva à influência dos seus pais que foram ambos professores do ensino primário. Inclusivamente a sua mãe foi sua professora no ensino primário e, além disso, teve a oportunidade de acompanhar, de uma maneira próxima, a profissão da sua mãe. Por exemplo, referiu que quando estava a realizar estágio profissional visitava a escola onde a sua mãe lecionava e assistia com alguma frequência às suas aulas. Estela recordou que no fim de uma destas aulas fez uma crítica à atuação da mãe, que fez com que esta alterasse o seu modo de interação com os alunos. De acordo com Estela, a sua mãe não evidenciava satisfação quando os alunos lhe davam as respostas desejadas às questões que lhes colocava. Também parecia ficar irritada com eles, quando lhe respondiam errado. Estela considerava que esses momentos referentes a duas situações distintas deveriam ser ponderados de um modo diferente pela sua mãe, revelando, de certo modo, uma certa sensibilidade para questionar e refletir sobre ações no ensino:

> Recordo-me perfeitamente de uma vez de lhe dizer assim… no fim de uma aula, 'Mamã, já reparaste que os alunos fazem bem e tu não fazes festa nenhuma, e quando eles fazem mal ficas tão zangada?! Como é que eles notam a diferença?'. Ela diz que nunca se esqueceu de eu ter dito aquilo e que mudou a partir daí. Porque ela, de facto, era assim. Embora notássemos que ela tinha orgulho de nós fazermos e querermos brilhar… na escola [primária] havia o costume de irmos às salas uns dos outros, ler as composições que era para as crianças desenvolverem a linguagem e o raciocínio (…) A partir daí ela mudou de atitude. (entrevista 1)

O pensar em ser professora de Matemática surgiu durante o 5.º ano, com uma nova professora que teve durante dois anos letivos consecutivos, no 2.º ciclo, que lhe apresentou a Matemática de uma forma organizada e coerente, tornando a disciplina acessível:

> No 5.º e no 6.º anos tive uma professora muito organizada, por isso achei que aquilo era lógico e fácil e gostei. Para mim… não tinha dificuldade nenhuma, e depois gostava de ensinar… via a minha mãe a explicar… e achava engraçado. (entrevista 1)

Além disso, Estela apreciava o facto de a Matemática ser uma disciplina que obrigava a algum treino de modo a ganhar experiência e saber lidar com todo o tipo de exercícios que poderiam surgir, por exemplo, nos momentos de avaliação escolar, e que lhe permitia ter um bom desempenho na disciplina. Ao explicitar esse gosto refere:

> Gostava de praticar, muito, fazia tudo… o que havia na altura. Na altura lembro-me que não havia livros, havia um manual e caderno de exercícios (…) que tinha exercícios resolvidos. Portanto, eu estava habituada a ver os exercícios resolvidos, depois resolvia sem olhar, e depois ganhava confiança para fazer os outros e davam sempre certo… achei que aquilo era fenomenal… era muito bom! (entrevista 1)

5.1.3 A escola e relação com colegas e alunos

No início da sua carreira docente, Estela lecionou durante cerca de seis anos consecutivos ao 3.º ciclo de uma mesma escola com ensino básico. Considera que o seu trabalho como professora se tornou mais desafiante e interessante quando tomou a decisão de mudar para uma outra escola com ensino secundário, especialmente, para ter oportunidade de lecionar os três níveis do secundário. Na sua opinião, com esta transição, passou a ter uma maior responsabilidade na gestão curricular da disciplina no sentido de promover no ensino uma melhor articulação e interligação entre os conteúdos matemáticos do ensino básico e do secundário:

> [Na escola básica] tinha sétimo, oitavo e nono, mas a maneira como se dá os conteúdos não é a mesma como se dá no sétimo, no oitavo e no nono nesta [na atual escola secundária com ensino básico]. Porquê?... Porque nós [na atual escola] como temos décimo, décimo primeiro e décimo segundo… sabemos perfeitamente o que é preciso no secundário do básico. Então na altura do básico nós já desenvolvemos alguns conteúdos que sabemos que vão ser mais necessários, em detrimento de outros que se dão no básico, enquanto que numa [escola] básica nós damos tudo como se tivesse tudo a mesma importância. De maneira que foi muito benéfico, embora o início tenha custado, ter vindo para uma escola secundária, a nível profissional. (entrevista 1)

Estela considera-se bem adaptada na atual escola secundária onde leciona, cujo ambiente considera, de uma maneira geral, muito bom para todos os intervenientes, em particular, para professores e alunos. Estela refere estar disponível para trabalhar com os colegas de outros grupos disciplinares sempre que lhe é solicitado. Por exemplo, recentemente, uma colega de *Física e Química* (entrevista 1) pediu-lhe que trabalhasse numa aula de apoio as conversões de unidades e a notação científica com os alunos de uma turma comum, porque esta colega tinha notado que estes estavam com algumas dificuldades nestes assuntos.

No que diz respeito ao grupo disciplinar, o trabalho que desenvolve com os colegas consiste na elaboração das planificações em grupo e na participação nas reuniões de grupo disciplinar ou de departamento. Este grupo habitualmente troca fichas de trabalho e, por vezes, alguns professores reúnem-se para elaborarem em conjunto fichas de trabalho e testes escritos. Na opinião da professora, na sua escola, há a prática comum dos professores apoiarem e participarem nas aulas uns dos outros, tratando-se de uma experiência que os próprios alunos não estranham:

> Nós aqui na escola estamos muito habituados e facilmente entramos nas salas de aulas uns dos outros… para ajudar… para colaborar. (…) Ou está mesmo no horário [a ida à sala de aula de um colega] e os alunos estão habituados… e reagem muito bem. Até com esta questão da avaliação se nós aparecemos na sala, é [os alunos acham que é] mais um [professor] como de costume. Ou então espontaneamente (…) marcamos (...) Estamos perfeitamente à vontade para trabalharmos em grupo. Mas isto claro… é uns mais do que outros … isto depende de cada um. (entrevista 1)

Estela sente que também pode partilhar as suas experiências profissionais com alguns colegas, por vezes, em conversas informais ou até por telefone ou por e-mail. Considera que faz parte da sua atitude profissional questionar os assuntos que leciona e sempre que tem ocasião tenta tirar as suas dúvidas de Matemática com os colegas da sua área na escola: "Tenho muita a mania de andar sempre a questionar os colegas… porque se me surge uma dúvida [sobre algum conteúdo]… tenho de perguntar a alguém para ver se tem a mesma ideia do que eu…pronto" (entrevista 1). Contudo, acha que são raros os professores de Matemática que ela conhece que questionem assuntos com colegas, particularmente na área de Estatística. Neste ano letivo teve a oportunidade de trocar algumas opiniões e de discutir alguns assuntos deste tema com uma colega que leciona a disciplina de Matemática Aplicada às Ciências Sociais no 10.º ano.

Estela alega que exerce a sua profissão com dedicação e satisfação pois, se assim não fosse, consideraria não ter condições para se manter nela. Considera que tem tido sucesso em transmitir o gosto que tem pela Matemática à generalidade dos alunos e pensa ter uma certa disposição para lidar com eles. Estes são também alguns fatores que favorecem a sua atitude positiva face à sua atividade profissional. Estela tem a convicção de ser importante haver na sala de aula um clima de boa disposição e de respeito mútuo e que, por essa razão, sempre que encontra algum aluno das suas turmas no recreio da escola, ou a entrar para a sala de aula, ou até mesmo fora do contexto escolar, faz questão de o cumprimentar. Na sua perspetiva, para se ter na sala de aula um ambiente de aprendizagem incentivador, no qual o respeito recíproco é essencial, este tem de ser trabalhado ou construído desde o início do ano letivo, com maior ou menor intensidade, dependendo das caraterísticas individuais dos alunos e globais da turma. Consequentemente, refere ter o cuidado de observar e tomar notas acerca dos comportamentos dos alunos na realização das atividades iniciais propostas, com a pretensão de identificar dificuldades em Matemática, nas interações entre os pares ou com ela, para posteriormente decidir como agir para os auxiliar.

5.1.4 Participação em projetos e formações

No contexto escolar, tem vindo a desenvolver alguns projetos com as suas turmas e normalmente participa nas várias atividades promovidas na escola, sejam do âmbito da Matemática ou de outras áreas, referindo alguns exemplos: o dia do Pi (a 3 de março); a receção na escola da exposição do *Atractor* da Universidade do Porto; a dinamização de atividades com frisos/padrões com uma turma do 8.º ano e de um *workshop* de jogos. Embora dependendo da oferta existente de formação para professores, Estela tem também participado em várias formações no âmbito da Matemática: "Participei em muitas e todas muito ligadas à área de Matemática, mas principalmente escolho Matemática. Não gosto de formações muito teóricas que não tenham nada a ver [com a prática]" (entrevista 1). Uma das primeiras formações que frequentou foi sobre fractais e depois outras se seguiram sobre tecnologias (calculadora gráfica, Excel, quadros interativos), sobre área escola e biblioteca, e também sobre tarefas matemáticas e tarefas estatísticas, esta última a mais recente.

A propósito da aplicabilidade que atribui às formações frequentadas, Estela comentou, por exemplo, que assim que terminou a formação sobre fractais, decidiu introduzir esse tópico ao lecionar o tema das sucessões numa turma de 11.º ano. Esta foi também uma forma de desenvolver mais este assunto, pois houve partes que precisou de estudar antes de as usar na sala de aula. De uma maneira geral, relativamente às formações, a professora diz habitualmente agir da mesma forma, tentando "…usar [na sala de aula] qualquer coisa que tenha lá feito, senão desinteresso-me pela formação" (entrevista 1). Além disso, participa regularmente nos *ProfMats*. Salientou que participar no *ProfMat* é sempre uma experiência enriquecedora; nas suas palavras: "é excecional (…) venho de lá estourada, mas com o coração cheio e com muita vontade de trabalhar, apesar de toda a burocracia [que se enfrenta atualmente nas escolas]" (entrevista 1).

5.1.5 A experiência pessoal com a Estatística

Relativamente aos temas que estudou enquanto aluna no ensino secundário, Estela não se recorda de ter aprendido Estatística. Relembra que apenas frequentou uma disciplina designada por *Probabilidades e Estatística*, durante o curso, no ensino superior, da qual recorda ter sido aprovada com uma classificação baixa, e de não ter entendido de uma forma aprofundada o que lhe foi ensinado. A esta distância, Estela tem dificuldade em recordar todos os temas tratados nessa disciplina mas afirma não ter aprendido a Estatística descritiva que leciona no ensino secundário e tem ideia de ter aprendido várias distribuições, tais como a distribuição de Poisson e a distribuição Normal, entre outras, e de usar tabelas estatísticas para resolver os exercícios:

> Lembro-me vagamente que não foi médias, nem desvios-padrão, nem nada disso… era a distribuição de *Poisson*, era a distribuição *Normal*… nunca percebi porque é que tinha de calcular a probabilidade de ser menor ou igual…, depois como eram tabelas… era só por observação de tabelas para calcular a probabilidade de estar ali, tinha de consultar as tabelas, se acertasse acertava, se não acertasse não acertava… porque eu não percebi. (entrevista 1)

Enquanto aluna de mestrado, na área de Educação Matemática, frequentou uma disciplina de introdução à Estatística. Recorda-se que o teste desta disciplina foi centrado nos assuntos teóricos abordados, que foram precisamente os assuntos que "não percebeu". Já para a dissertação, lembra-se de ter estudado tópicos "mais práticos", nomeadamente, alguns tipos de amostragem tendo desenvolvido pouco a parte dos testes de hipóteses estatísticos, tal como refere neste excerto:

> Engraçado, que fui estudar [no mestrado] uma coisa que agora dou no ensino secundário em Matemática Aplicada às Ciências Sociais… [estudei] mas foi tipos de amostragem e falei pouquinho em testes de hipóteses (…) gostaria de estudar mais sobre testes de hipóteses. (entrevista 1)

Estela considera que, se alguma vez tiver oportunidade de voltar a estudar, desejaria optar por uma área que lhe permitisse desenvolver a capacidade de interpretação de problemas. É da opinião que uma experiência deste género lhe poderia ser bastante útil para a docência, em particular, para ensinar o tema da Estatística. Contudo, fez também notar que, no ensino secundário, essa vertente interpretativa é mais explorada na

Estatística lecionada na Matemática Aplicada às Ciências Sociais do que na Matemática A:

> Tenho dificuldade na interpretação do problema, interpretação dos dados, interpretação do gráfico, gostava… Tudo tem a ver com a experiência no básico, que tivemos como aluno… e, como professora, eu gostava de desenvolver isso… e geralmente, na Matemática A não há [interpretação]… Onde é que tive mais gosto? [no sentido de desenvolver um pouco mais essa capacidade]... Tive mais gosto em MACS (…). (entrevista 1)

Na entrevista inicial, acaba por destacar alguns exemplos de conhecimentos estatísticos que adquiriu enquanto aluna de mestrado, e que lhe foram úteis para o ensino da Estatística. Por exemplo, Estela salientou a aplicação dos conhecimentos sobre amostragem nas suas turmas de MACS e referiu, adicionalmente, ter sido também alertada sobre o cálculo do desvio-padrão que é diferente para o caso dos dados provenientes de uma amostra ou para o caso em que se possui todos os dados de uma população. Aos referir-se a essa distinção, menciona ainda: "Quem eram os professores que estavam atentos a isso? (…) Os manuais não faziam isso!" (entrevista 1). Na sua opinião vários conhecimentos estatísticos não eram do conhecimento geral dos professores.

Da experiência como formanda numa ação de formação no âmbito da Estatística que teve como formadora a investigadora do presente trabalho, Estela pondera a possibilidade de aproveitar algumas ideias presentes nas tarefas desenvolvidas nessa ação para a sua prática, em particular, "o relacionar representações (gráficas e/ou numéricas) para desenvolver mais os conceitos. (…) Vai-me dar muito mais trabalho, mas é capaz de me dar muito mais gosto preparar este tema [desta vez]" (entrevista 1). Ao refletir sobre os conhecimentos que possui de Estatística, nomeadamente, os que deve ensinar no 3.º ciclo e no ensino secundário, Estela constata que estes foram adquiridos sobretudo no decurso da sua prática profissional, na preparação que tem feito para lecionar este tema aos seus alunos, com o apoio fundamental de vários manuais escolares.

5.1.6 A turma e o contexto letivo

A turma onde foi desenvolvido o presente estudo tem 25 alunos e é composta por 13 alunos do sexo feminino e 12 do sexo masculino. Estela considera que se trata de uma turma com um bom comportamento. Em termos de classificações e do desempenho da globalidade dos alunos, a professora considera tratar-se de uma boa turma, uma das melhores turmas de 10.º ano da escola, pois, no final do segundo período, os alunos, de uma maneira geral, tinham alcançado boas notas nas diversas disciplinas e havia poucos alunos com classificações negativas.

A escola onde Estela leciona está totalmente renovada. Tal situação reflete-se nas caraterísticas das salas de aulas, ou seja, são espaços com novos equipamentos, constituídos por mesas para dois alunos, com projetor de dados, quadro interativo e computador na secretária da professora. Todas as aulas assistidas foram realizadas em salas similares.

O ambiente nas aulas de Estela pode caraterizar-se como informal, com um bom ritmo de trabalho desde o início ao fim da aula. A aula funcionava com participação dos alunos na resolução das tarefas propostas e também com um grande envolvimento da professora, em particular, na discussão e correção da maioria das tarefas trabalhadas na aula, embora sempre em interação com os alunos, dando alguma oportunidade para que estes formulassem e esclarecessem as suas dúvidas. De uma maneira geral, as aulas de Estela estavam estruturadas e organizadas em torno da resolução de tarefas que, na maioria das vezes, eram trabalhadas pelos alunos com o apoio da calculadora gráfica. Uma estratégia de ensino habitual que Estela usava era a introdução de conceitos ou a revisão de algum conteúdo quando os alunos se encontravam a trabalhar numa tarefa que abrangesse esses assuntos.

O ambiente geral era também de um clima de boa disposição e havia algum ruído nos momentos em que os alunos trabalhavam nas tarefas aos pares ou em pequeno grupo com os colegas que estavam em mesas contíguas, provocado sobretudo pela troca de impressões. Nestes períodos, a professora circulava entre os grupos, observando-os, apoiando o trabalho deles, e geria o tempo. Os alunos pareciam concentrados nos seus trabalhos. Por vezes, a professora tinha necessidade de alertar alunos que pareciam estar distraídos com outros assuntos para estarem atentos à aula, agindo sempre de uma forma humorística e os alunos visados reagiam bem e alteravam os seus comportamentos. Os alunos não pareciam ter estranhado a presença da investigadora no decurso da observação de aulas. Vários alunos solicitaram-lhe que se aproximasse e ajudasse a superar dúvidas. Estela não esperava que os alunos reagissem de outro modo à presença da investigadora, pois considera que eles já estavam habituados a ter nas aulas a presença de mais do que um professor, nomeadamente, da área da Matemática, por esta ser uma prática frequente do grupo de professores de Matemática da escola.

Estela procurava envolver os alunos na aula, quer nos momentos de exposição dos conteúdos que fazia, quer nos de correção dos trabalhos. Solicitava-lhes que respondessem a questões e também, por vezes, requeria-lhes uma ida ao quadro ou ao computador (cujo ecrã estava usualmente projetado na tela) para mostrar à turma os passos da resolução de alguma questão na calculadora gráfica, mesmo que estes alunos dissessem que não estavam completamente à vontade no manuseamento da calculadora ou com algumas dúvidas na tarefa proposta em mãos.

Na turma, todos os alunos possuíam calculadora gráfica. A maioria dos alunos usava a máquina da marca *Texas Instruments (TI 84/Plus)*, havendo apenas dois alunos que possuíam calculadoras gráficas da marca *Casio*. Dos alunos que usavam a referida *Texas Instruments* (TI), quatro alunos possuíam e traziam adicionalmente a calculadora gráfica avançada *TI-Nspire* da *Texas Instruments*.

No ensino do tema da Estatística, Estela usou sempre ou quase sempre a calculadora gráfica, quer a física (da marca *TI*) quer a que possui instalada no computador (também da marca *TI*) para ser usada através da sua projeção na tela da sala de aula. Esta projeção era realizada sempre que necessitava fazer um esclarecimento acerca do uso da calculadora ou quando pretendia envolver alunos para mostrarem os procedimentos úteis que tinham descoberto na exploração da calculadora habitual ou da mais avançada, sempre num contexto de exploração e resolução de tarefas.

5.1.7 Síntese

Estela é uma professora com experiência no ensino secundário, com uma postura inquisitiva sobre os assuntos que leciona, o que a conduz muitas vezes a procurar a opinião dos colegas do grupo disciplinar quando tem alguma questão ou dúvida sobre algum conteúdo que tenha de ensinar. Prestável, atenciosa e dinâmica, mantém uma atitude aberta e disponível para participar em novas iniciativas, para colaborar com os seus colegas na escola e para participar em formações no âmbito da Matemática.

Relativamente aos seus conhecimentos sobre Estatística, pensa que estes foram principalmente adquiridos na sua prática profissional apesar de ter frequentado durante o curso, no ensino superior, uma disciplina designada por *Probabilidades e Estatística*, na qual não se recorda de ter aprendido Estatística descritiva e o processo de amostragem que ensina no ensino secundário. Mais tarde, durante o mestrado teve oportunidade de explorar alguns conceitos estatísticos e a amostragem; contudo, sente que tem ainda muito para aprender sobre Estatística, referindo-se, especificamente, ao processo de amostragem e à necessidade de ganhar experiência na interpretação dos problemas. Por conseguinte, manifesta desejo de aprofundar esses assuntos, pensando fundamentalmente na contribuição que poderia dar para o desenvolvimento do ensino da Estatística no ensino secundário na disciplina de Matemática Aplicada às Ciências Sociais que, na sua opinião, tem um maior enfoque nesses conhecimentos. Da sua experiência na participação na ação de formação que serviu de contexto para a seleção de participantes para o presente estudo, conta levar para a sua prática situações que incidam em estabelecer conexões entre representações para promover o entendimento de alguns conceitos estatísticos nos alunos.

Estela destaca que a turma onde decorre este estudo é uma das melhores da escola e que os seus alunos são, de um modo geral, trabalhadores e responsáveis. A boa disposição e bom ritmo de trabalho espelham, de uma maneira geral, o ambiente das aulas. O trabalho desenvolvido nestas aulas é focado na resolução de tarefas com o apoio da calculadora gráfica. A professora tem um papel central na aula, na sua gestão, na discussão e na correção das tarefas, entre outras, mas também tem o cuidado de incentivar a participação dos alunos nas atividades propostas.

5.2 Perspetiva geral sobre o conhecimento didático do professor em Estatística

5.2.1 O programa da disciplina de Matemática A

Ao situar a Estatística nos diferentes programas escolares de Matemática A ao longo da sua carreira docente, Estela relembrou que chegou a lecionar Estatística no ensino secundário ainda de acordo com o programa de Matemática resultante do ajustamento dos programas do ensino secundário (ME, 1997). Estela é da opinião que o tópico das variáveis bidimensionais era o conteúdo que este programa, na altura, trazia de novo e crê que o programa vigente[15], homologado em 2001, mantém os mesmos conteúdos do programa de 1997. Estela considera que o diagrama de extremos e quartis, as medidas de dispersão e as distribuições bidimensionais ou bivariadas são os assuntos do programa do secundário que a maioria dos seus alunos vai contactar pela primeira vez.

Durante a entrevista inicial, a investigadora pediu a Estela para realizar uma leitura do programa de Matemática A vigente (ME, 2001) focada essencialmente nos conteúdos programáticos a abordar e guiada pelo quadro do tema Estatística incluído no programa (anexo 5), que a investigadora lhe colocou à disposição. Este quadro destacava os três tópicos principais do tema, nomeadamente: *Generalidades sobre Estatística*; *Organização e interpretação de carateres estatísticos* e *Referência a distribuições bidimensionais* (no mesmo quadro também havia referência a subtópicos e indicações metodológicas gerais de cada tópico). Houve o seguinte diálogo:

> **Investigadora:** Qual a tua opinião sobre os conteúdos do programa?
> **Professora:** Estou a ver… Ora [diz aqui] 15 aulas, portanto são 5 semanas…
> **Investigadora:** Sim, cerca de 5 semanas…além da tua opinião também queria saber como os desenvolves…
> **Professora:** (lê o quadro alto e bom som) Lá está!... População e amostra [os alunos] tem uma noção intuitiva [destes conceitos]… recenseamento… está [agora] a decorrer um recenseamento… sondagens [os alunos] estão fartos de ouvir falar por causa das eleições na comunicação social. Mas... por exemplo… uma coisa que me intriga de dar é quando é que eu tenho a certeza que uma amostra é representativa da população? Qual é que deve ser? Ainda há pouco tempo vi num livro que uma população considerava-se infinita quando for, acho eu que era, … se o seu tamanho fosse maior ou igual… não, não… era maior do que vinte vezes a dimensão da amostra que escolher.

[15] Ao longo de todo este capítulo, as referências ao programa de Matemática A do ensino secundário dizem respeito ao programa em vigor no momento da recolha de dados (ME, 2001).

> **Investigadora:** Há algumas técnicas que permitem obter estimativas para as dimensões das amostras…
> **Professora:** Exatamente, mas como é que eu sei [as que interessam]?… Agora os gráficos de barras, gráficos circulares [os alunos] já sabem do 8.º ano… os pictogramas… já não se usam… muitas vezes, não surgem nos manuais… são de uma falta de rigor completa! Mas eles [os alunos] até sabem criticar. A moda é fácil. Depois…nos gráficos… a função cumulativa no caso da variável discreta quase que nem se usa… no caso contínuo serve para localizar quartis e mediana graficamente… com frequências relativas… essencialmente. As limitações das estatísticas [medidas de localização e dispersão], diagramas de extremos e quartis… Mas isso tudo (referindo-se aos tópicos e subtópicos do programa) é Estatística descritiva! O coeficiente de correlação e a reta de regressão calculam-se até na calculadora gráfica. Dá-se o centro de gravidade e a reta de regressão também na calculadora e mais nada!… Não está aqui Estatística indutiva quase nenhuma. Não está aqui a generalização à população, não está! (…) na Estatística da Matemática A é tudo muito descritivo, faz-se cálculos e gráficos na calculadora, não há, de uma maneira geral, necessidade de se recorrer à interpretação. (…) Os alunos não estão habituados a discutir conceitos estatísticos nem resultados. (entrevista 1).

Estela dá a entender que vários conceitos e representações gráficas referidos no tema da Estatística do programa de Matemática A são reconhecidos pelos alunos, uns porque já lidaram com eles no 8.º ano (gráficos de barras, gráficos circulares), outros através da experiência do dia-a-dia, nomeadamente por meio da comunicação social (recenseamento, sondagem). Considera que o pictograma, e mesmo a função cumulativa, têm uma aplicabilidade limitada, o primeiro devido à sua falta de precisão e a segunda usada para identificar quartis de uma distribuição com dados agregados. A professora também aponta um conjunto de questões relacionadas com amostra sobre as quais gostaria de estar mais esclarecida, nomeadamente, quando é que uma amostra é representativa, definições de população finita e infinita e dimensão de uma amostra.

Quanto ao estudo das distribuições bivariadas menciona o papel central da calculadora gráfica neste estudo, que permite obter reta de regressão, coeficiente de correlação e centro de gravidade. Estela associa o tema da Estatística do programa à Estatística descritiva, que tem um enfoque no cálculo e nas representações, e no qual a interpretação dos resultados não é privilegiada, considerando mesmo que a presença da Estatística indutiva ou inferencial é praticamente inexistente.

Quando questionada sobre a realização de estudos estatísticos que envolvam as quatro fases do processo de investigação estatística, a professora considera que este não é explicitamente enfatizado no programa de Matemática A; contudo, referiu também que, na sua prática, tem vindo a integrar a realização de algum tipo de estudo estatístico. Isto tem acontecido através da troca de experiências profissionais com colegas que lecionam os mesmos níveis e das pesquisas que tem feito na internet à procura de tarefas para propor na sala de aula. Por exemplo, Estela contou que, na sua prática, no ano letivo anterior, propôs uma investigação estatística, no módulo da Estatística, à sua turma do secundário do ensino profissional. Pretendia que os alunos fizessem "um estudo sobre a obesidade de todos os alunos de 10.º ano da escola" (entrevista 2). Achava que este trabalho os "iria motivar" por ser um tipo de trabalho que os próprios alunos nunca tinham feito e, por outro lado, poderia ser que descobrissem "informações interessantes" acerca da situação em estudo (entrevista 2). Ao longo deste processo surgiram várias contrariedades: a recolha de dados que foi feita pelos alunos desta turma do secundário tornou-se muito demorada (havia oito turmas para recolher dados); os alunos revelaram dificuldades em trabalhar com o *Excel* (por conseguinte, foram dadas algumas aulas na biblioteca da escola, focadas em funções estatísticas que pudessem ser úteis para a análise de dados que os alunos teriam que fazer); e, entretanto, foi atingido o limite de aulas do módulo da Estatística. Estela teve então de dar seguimento a outro tema, e o estudo que estavam a fazer acabou por ser desenvolvido pelos alunos, em grupos, mas fora do contexto de sala de aula. Os trabalhos finais foram enviados para o e-mail da professora: "… queria habituá-los e responsabilizá-los a mandar-me as coisas dentro de um determinado prazo… mas não tiraram grandes conclusões" (entrevista 2). Apesar do esforço que foi feito para superarem os obstáculos que foram surgindo ao longo da realização desta investigação, pelo que se lembrava, Estela descreveu os relatórios finais da seguinte forma:

> Eles… continham sobretudo gráficos…. Houve um grupo que fez o gráfico circular … outros gráficos… e escreveram no relatório o que achavam da sua turma relativamente a uma outra [que escolheram]… mas a maior parte não fez isso… (entrevista 2)

Para Estela, os resultados não foram muito surpreendentes, pois os alunos tinham que lidar com muita informação e, contrariamente ao planeado, não pôde apoiá-los na análise dos dados em aula. Além disso, foi a primeira vez que estes alunos do ensino profissional realizaram um trabalho desta natureza e tiveram de usar um recurso sobre o qual não tinham grande experiência. A professora não ficou admirada com o facto dos trabalhos finais dos alunos exibirem sobretudo representações gráficas, havendo apenas um grupo que comparou dados recolhidos de duas turmas.

A professora tem a perspetiva de que uma atividade dessa natureza não faria sentido na disciplina de Matemática A; contudo, caso tivesse tempo para propor na aula uma atividade de cariz investigativo, via-se a fornecer "a questão e dados" e a sugerir aos alunos para "trabalharem a análise dos dados e a apresentação dos resultados" (entrevista 1).

Por várias vezes, nas diferentes entrevistas, Estela refere a necessidade do programa de Matemática A, no tema da Estatística, ser mais desenvolvido ao nível dos conceitos e incluir partes da Estatística inferencial. Aponta como uma das razões o facto deste tema ser apresentado com maior profundidade no então novo programa do primeiro ciclo (ME, 2007), o que deveria naturalmente "implicar mudança" no mesmo tema no programa do ensino secundário. Outra razão que indica para uma possível mudança é considerar que os alunos do ensino secundário, como referido anteriormente, deveriam ter oportunidade de experienciar a "Estatística inferencial" que está mais relacionada com "questões reais" (entrevista 1). Neste sentido, considera que, no ensino secundário, poderia ser útil:

> Apresentar-se mais exemplos práticos e discutir [possíveis] soluções… Devia-se estudar uma amostra e as conclusões que se podem inferir para a população… A Estatística inferencial, a generalização à população não está aqui! [apontando novamente para o resumo do programa] (…) Por isso é que os alunos me dizem "Oh, stora… estou mortinho para chegar à Estatística… quando é que vamos iniciar a Estatística?"… Porque eles já sabem a maior parte [dos conteúdos]… (entrevista 1)

A professora manifesta desta forma a perspetiva que tem da Estatística escolar e que esta deveria ser estudada de uma forma mais aprofundada ao nível da Matemática A, no ensino secundário.

Estela é também da opinião que a generalidade dos alunos de Matemática A, no fim do estudo da Estatística no 10.º ano, retém a ideia que o tema é acessível "Deve ser o tema que mais gostam (…) É fácil… depois as tarefas que fazem dá-lhes sempre certo!" (entrevista 1), mas "não é para aplicar no dia-a-dia" (entrevista 1), no sentido de que não vê estes alunos a ter que recorrer "à elaboração de um gráfico" (entrevista 1) para tomar decisões para a sua vida pessoal. Na sua perspetiva, neste nível de ensino e tema, os assuntos mais desenvolvidos acabam por ser precisamente a construção ou determinação de representações gráficas (gráficos de barras, diagrama de extremos e quartis, diagrama de dispersão, reta de regressão, etc.), que são conteúdos que não vão ser colocados em prática pelos alunos nas suas vidas diárias.

5.2.2 Planificação da unidade de Estatística

Estela crê que, atualmente, de uma maneira geral, atua do mesmo modo, quer na preparação do tema da Estatística, quer dos restantes temas lecionados no 10.º ano. Habitualmente seleciona tarefas de manuais e dos cadernos de apoio ao professor, prepara fichas de trabalho e acetatos, dá uso ao manual escolar e aos CDs que integram os diversos manuais, e recorre também às tecnologias (computador e calculadora gráfica) na resolução de tarefas, utilizando, também, por vezes, *software* multimédia (por exemplo, *e-escola* – Porto Editora) para apresentação de conteúdos e exemplos diversos nas aulas. Estela transmitiu ao longo da entrevista inicial a ideia que a sua abordagem é semelhante em todos os temas do programa: "[Na Estatística] não uso nada a mais do que nos outros temas… em qualquer tema se é atrativa determinada atividade… faço-a" (entrevista 1).

Estela admite que utiliza bastante a calculadora gráfica em qualquer tema, mas possivelmente com maior frequência neste tema: "se calhar, uso-a mais na Estatística" (entrevista 1). Na opinião da professora, a calculadora gráfica, além de ser um recurso de uso obrigatório pelo programa, é um recurso cuja utilização é muitas vezes solicitada nos testes e exames nacionais que os alunos realizam no ensino secundário, o que implica que estes devem aprender a dominar não só a vertente do cálculo como a vertente gráfica na sua utilização. Estela deu como exemplo o caso da escolha da janela de visualização adequada, situação em que os alunos muitas vezes apresentam dificuldades no tema das funções:

> [os alunos] têm de comprar [a calculadora] porque têm exame de 12.º ano e nos testes intermédios há sempre perguntas que envolvem a calculadora gráfica (…) cheguei a dar aos meus de 10.º ano nas funções uma pergunta que dizia 'utilizando as capacidades da tua calculadora gráfica, calcula os zeros e decompõe ou fatoriza o polinómio…'. Uma coisa é usar sem recorrer à calculadora gráfica e outra coisa é usando as capacidades da calculadora gráfica… claro que são exercícios que não é só meter, por exemplo, a função e

> ver-se logo os zeros… não… são exercícios que implicam a escolha de uma janela de visualização que muitas vezes eles têm muitas dificuldades nisso. (entrevista 2)

Para Estela, essa dificuldade na escolha da janela de visualização é também evidente na Estatística. Ainda neste tema, a professora expressou a utilidade de uso da calculadora gráfica quando resolvem uma tarefa com vários dados, libertando os alunos do peso dos cálculos: "Porque agora [no ensino secundário] não são obrigados a saber calcular aquelas fórmulas com muitos dados, calcular a média é impensável… calcular a mediana é impensável… os quartis, o desvio-padrão (…)" (entrevista 2). Na sua perspetiva, a calculadora gráfica é particularmente adequada para se trabalhar com um maior número de dados de modo mais eficiente e para se obterem representações que podem contribuir para a realização de análises mais completas.

Estela refere que tenta diversificar o modo como leciona o tema de ano para ano. Usualmente, não consegue seguir as planificações e usar materiais de aulas de anos letivos anteriores sobre os mesmos assuntos e tem necessidade de as modificar para levar à sala de aula algo de novo de modo a proporcionar um ensino mais estimulante para os alunos e também para si própria. Estela assume que nem sempre atuou da mesma forma na preparação dos temas a lecionar, pois, ao recordar a primeira vez que lecionou ao 10.º ano Matemática A, e referindo-se particularmente aos materiais que preparou e levou para a sala de aula, destaca a discrepância entre a quantidade de materiais que usou no tema da Geometria e os que usou no tema da Estatística; neste último foi muito reduzida, condicionada por não dominar da mesma forma os conhecimentos que possuía destes temas:

> … por exemplo, lembro-me na primeira vez que dei ao 10.º ano, na Geometria trouxe o nível de bolhas de ar, trouxe o fio de prumo. O meu pai andou a comprar-me o nível de bolhas de ar, que eu nem sabia o que era aquilo, e andei a experimentar como é que aquilo funcionava em casa… por causa dos critérios ao nível de paralelismo e perpendicularidade… depois com os esquadros a pôr uma antena direita no telhado, quando é que a reta era perpendicular ao plano. Trazíamos cubos, em plasticina, cubos em acrílico, com líquidos. Enquanto que para as outras matérias da Matemática nós trazíamos montes de coisas, para a Estatística o máximo que trazíamos, nem sequer era o computador, eram os acetatos com as coisas direitinho… a dar tudo certo. Porque as coisas não se dominavam. (entrevista 1)

Estela também indicou que, na primeira vez que lecionou o tema da Estatística, a sua falta de experiência no ensino deste tema e o pouco tempo que teve para o lecionar condicionaram fortemente a variedade de propostas de trabalho que levou para a sala de aula. Essas ficaram restritas a uma seleção de tarefas de alguns manuais, mas sobretudo do manual adotado.

Atualmente, Estela admite que na planificação e preparação das aulas de Matemática de qualquer tema tem por hábito desenvolver um trabalho cuidadoso à volta das tarefas a propor na sala de aula, referindo-se concretamente à sua escolha, adaptação ou criação. Estela refere usar muitas vezes os cadernos de apoio ao professor para selecionar tarefas para desenvolver nas aulas em que vai lecionar o tema da Estatística. Disse que, frequentemente, não tem disponibilidade suficiente para proceder a uma análise mais cuidada das tarefas ou atividades destes cadernos de apoio ao professor, que habitualmente vêm com muita informação:

> (…) para ler isso tudo [apontando para uma tarefa de Estatística de um caderno de apoio]… (...) e decidir qual a que me interessa ou qual a que resultará para os meus alunos, não é fácil! (…) Depende do tempo que tiver. (entrevista 1)

Contudo, considera que tem sempre uma tarefa introdutória para um determinado conteúdo que pode selecionar dos cadernos de apoio ao professor ou mesmo de manuais, mas dá preferência às situações provenientes de cadernos ou manuais que não sejam do adotado. E para exemplificar um tipo de tarefa introdutória na Estatística a professora apresenta o seguinte:

> [no tema das Funções] quando iniciei a função quadrática, dei uma ficha de trabalho que direcionava… [algo do género] "introduza na calculadora gráfica… cada uma destas situações… agora veja o que acontece e descreva cada uma delas". Na Estatística posso fazer uma atividade semelhante, posso pedir para que se desenhe a nuvem de pontos, por exemplo,… os alunos têm que saber dominar a calculadora, não é? (entrevista 1)

Estela assume que o manual adotado tem um papel de destaque na atividade de consolidação dos conhecimentos estatísticos dos alunos. Além disso, também considera que os manuais atuais, comparativamente com os de uma década atrás, apresentam alguma evolução nas tarefas que reúnem, que lhe parecem mais focadas em situações reais do que fictícias, as quais, na sua opinião, acabam por suscitar maior interesse nos alunos.

A professora disse ainda que tem por prática geral produzir fichas de trabalho nas quais inclui essas tarefas, que distribui nas aulas para serem trabalhadas pelos alunos. Contudo, nem todas essas tarefas são para resolver na aula, embora devam ser resolvidas em casa, tal como refere: "Mas isso já vai do hábito que se criou desde o início do ano. Mando uma ficha de trabalho [por e-mail] e não preciso de lhes dizer quais as tarefas que têm de fazer" (entrevista 2). Para Estela, a elaboração destas fichas de trabalho para qualquer tema é também uma forma de ajudar os alunos a focarem-se mais no leque diversificado de tarefas com maior probabilidade de surgirem no exame final da disciplina.

5.2.3 Caraterísticas das tarefas estatísticas

Para Estela uma tarefa pode ser um exercício ou pode ser uma ficha de trabalho centrada numa situação ou conteúdo específico. Ao explicitar melhor o que entende por tarefa disse:

> Porque no fundo com uma tarefa estou a questionar e relacionar uma série de coisas. Acho eu, que o objetivo de uma tarefa é esse, não é?... Acho que é pegar numa coisa que os obrigue a pensar… a procurar… a pesquisar… depois a explicar… Pronto. Depois dizemos que está certo ou não. (entrevista 2)

Neste excerto, Estela expressa que uma tarefa pode conduzir os alunos à exploração e explicação dos assuntos envolvidos. E fornece também uma indicação do papel que reserva a si mesma na avaliação e confirmação do trabalho desenvolvido.

Um tipo de tarefa muito frequente na sua aula são os exercícios que são geralmente compostos por questões "direcionadas" (entrevista 2). Mas Estela também afirmou que estes podem incluir questões "exploratórias" (entrevista 2) que induzam ir além da realização de cálculos e da leitura de dados numa representação gráfica:

> Por exemplo, tirei de livros [manuais] exercícios que não eram do tipo calcula a média, ou as medidas de tendência central… não era isso… era olhar para um gráfico e já ter uma visão crítica daquilo [referindo-se, por exemplo, à tarefa

> *Visitas a um museu* (anexo 2)]… para mim, exercício é tudo isso, utilizo… não tive nenhuma formação ainda para isso, mas vou puxando pela cabeça… vou observando… vou vendo. (entrevista 2)

Por motivo de considerar o ensino secundário mais exigente para os alunos, Estela confessa que nestes níveis recorre frequentemente às fichas de trabalho. Às vezes, usa-as para alguma atividade na aula, outras vezes, elas servem de complemento dos próprios manuais, uma vez que reúnem um conjunto de tarefas para os alunos praticarem em casa. Tem por hábito mandá-las pela internet para os e-mails dos alunos e todas as tarefas destas fichas que não são feitas nas aulas podem ser discutidas com ela nos intervalos ou na aula de apoio semanal que dá: "Na aula de apoio… quem quisesse tirava dúvidas… ou no intervalo … eles estavam à vontade. Têm de se habituar a trabalhar" (entrevista 2).

Estela elaborou 47 fichas de trabalho, dos temas da Geometria e das Funções, desde o início do ano letivo até ao momento em que iniciou o tema da Estatística. Para este tema, elaborou apenas quatro fichas de trabalho. Ao comentar sobre a experiência que teve na construção destas fichas de trabalho para o atual ano letivo, Estela revelou ter-se deparado com um manual não totalmente satisfatório ao nível das tarefas no tema das Funções, embora este lhe parecesse adequado no tema da Estatística. Considera que o manual, no tema das Funções, incorpora várias tarefas que não exploram os conteúdos do currículo da forma mais adequada e que estas não seriam totalmente benéficas para as aprendizagens presentes e consequentes do tema, no ensino secundário, embora as considerasse passíveis de resolução. Deste modo, a professora sentiu-se obrigada a produzir mais fichas de trabalho sobre esta temática:

> Por acaso, este ano não gostei do manual [sobre Funções] e havia tarefas que não interessavam a ninguém… não é que não fossem possíveis de serem feitas… não são as mais habitualmente exploradas. Com o tempo que se tinha… só os ia distrair daquilo que realmente interessava para depois para um 11.º ano, principalmente nas Funções … Do volume da Estatística gostei muito, mas o das Funções não! (entrevista 2)

No momento da primeira entrevista, Estela ainda não tinha iniciado o ensino do tema da Estatística, embora já tivesse feito as primeiras pesquisas ao nível das tarefas, principalmente no manual adotado. Afirmou que pretende sempre que as tarefas sobre o tema da Estatística, na disciplina de Matemática A, incluam questões diretas, de natureza semelhante às que propõe habitualmente na Matemática, em que o pedido de interpretação de algum resultado ou discussão de uma situação não é o mais comum: "as questões têm de ser direcionadas para responder àquilo [algo concreto]. A Matemática é só aquilo, não é para discutir, desde quando!? … é exata…e acabou!... na Estatística é igual (…) [as tarefas são] com pouco foco na interpretação. (entrevista 1)

Neste sentido, Estela também assume que as tarefas que propõe na Estatística são semelhantes às dos outros temas lecionados na Matemática. Por várias vezes nas entrevistas e mesmo nas aulas observadas, Estela menciona que na formação que frequentou sobre Estatística trabalhou com um número muito grande de dados nalguns problemas, algo a que não estava habituada. E mais especificamente para a sua prática com os alunos de Matemática A que tem, Estela considera que esta situação não é adequada nos moldes em que o programa está feito. E acrescenta:

> Achas que um professor se arrisca [a trabalhar com tantos dados]? (…) Se calhar uma Estatística que se pretenda crítica, aplicável, envolvendo [questões] do dia a dia (…) tem de ser um trabalho contínuo desde o básico ao secundário. Para que ao apresentarmos uma tabela, com pelo menos estes dados [apontando para uma tabela de uma tarefa do manual] saibam ler e ver o que mais interessa. Isto não é fácil! (…). (entrevista 1)

Contudo, Estela considera que na Estatística de MACS, habitualmente, há mais tempo para desenvolver o tema e pode-se criar oportunidades para envolver os alunos a trabalhar com problemas estatísticos com um conjunto de dados reais mais alargado. O que não deixa de ser um trabalho desafiante e difícil, sobretudo para um professor que não tenha muito à vontade e experiência em lidar com a amostragem, por exemplo, e com a interpretação e comunicação dos resultados:

> Em MACS, na próxima vez [que for a lecionar], por exemplo, na inferência estatística… para abordar o teorema do limite central… se calhar já vou usar estes dados todos que estão aqui [lista com peso de cada aluno de 10.º ano da escola onde leciona], estes são dados reais… que são da nossa escola, são de todos os alunos de 10.º ano, separados por sexo…E agora imagina dar isto a uma turma de Matemática A de 10.º ano?! [não há tempo] E agora trabalhar isso?... E depois gerar amostras…de dimensão 50… até eu terei de aprender a interpretar isso. Não é fácil… (entrevista 1)

Estela acha que a calculadora gráfica é uma ferramenta útil para ajudar os alunos a lidar com os dados e os cálculos de uma forma mais expedita e considera que esta, nas mãos de um professor, tem a potencialidade de contribuir para o desenvolvimento do seu conhecimento estatístico:

> E até nós professores que não tivemos essa formação inicial [na área da Estatística]… a calculadora gráfica serve para nós aprendermos a desenvolver mais algumas ideias e mais capacidades sobre Estatística… pode gerar mais discussão, dá para concluir mais coisas do que coisas evidentes. (entrevista 2)

Estas perspetivas da professora advêm do facto de achar que desenvolver um ensino da Estatística com um maior enfoque nos problemas e sua discussão, implicaria um maior domínio de conhecimentos estatísticos e que se proporcionassem aos alunos algumas experiências estatísticas desde o ensino básico para que não estranhassem totalmente a atividade a desenvolver.

Nessa ocasião, Estela indicou algumas tarefas do manual adotado que poderia propor nas aulas. Eram tarefas efetivamente "orientadas", nas palavras da professora, que envolviam alguma representação gráfica e questões específicas relativamente a essa representação e ainda alguns cálculos. Quando a investigadora lhe apresentou as duas tarefas *Média sobre média* e *Tempo de espera pelo autocarro* (anexo 2) e a questionou sobre a viabilidade de estas duas tarefas poderem ser propostas na aula, Estela disse, depois de as analisar, que poderia considerar que estas iam ao encontro das tarefas que habitualmente propõe no ensino do tema da Estatística. Fez apenas um reparo à tarefa *Tempo de espera pelo autocarro* que tinha a singularidade de, além de pedir o resultado do desvio-padrão, pedir também a interpretação desse resultado, salientando: "Isto não é costume interpretar!... Vou fazer isto porque vou ouvindo, com os colegas da formação… agora tenho mais sensibilidade" (entrevista 1). Ou seja, a professora reconhece que as tarefas que costuma selecionar sobre o desvio-padrão habitualmente não incluem perguntas que peçam a interpretação dos valores envolvidos.

A professora revelou ter noção do maior grau de desafio que tarefas exploratórias com vários dados reais possam implicar em termos de orientação do trabalho dos alunos e em se conseguir dar resposta às questões dos alunos. Por exemplo, quando a investigadora lhe mostrou a tarefa *Como está a evoluir a nossa população?* (anexo 2), apesar de considerar que continha poucos dados, Estela achou que os alunos iriam ter muitas dificuldades em perceber qualquer uma das primeiras quatro orientações de trabalho ali incluídas. A professora confessou ter, de uma maneira geral, uma expectativa muito baixa no desempenho dos alunos na sua resolução caso a propusesse na aula. Disse tratar-se de uma tarefa "aberta" devido à sua natureza exploratória e com potencial para originar "discussão". Para Estela os alunos estão habituados a que ao nível das orientações e questões "A primeira é iniciação, depois é mais um bocadinho que a outra. Não entendem que [no caso desta tarefa] está tudo a par" (entrevista 1), ou seja, o grau de dificuldade das questões que habitualmente coloca aumenta gradualmente e não lhes coloca um conjunto de sugestões passíveis de serem usadas ou não. Além disso, nas referidas orientações da tarefa há conceitos que a professora considera que os alunos não saberiam como utilizar na resolução da tarefa, destacando: "*Procure padrões?*, *Variabilidade dos dados?*, *Conjeture sobre os dados e seu contexto*? [ao ler da lista de sugestões da tarefa]", deixando transparecer que os alunos desconhecem os seus significados. A única parte das orientações da tarefa que Estela considerou, nas suas palavras, "a mais direcionada" e em consonância com os aspetos que trabalha na sala de aula era a última que indicava algumas representações numéricas ou gráficas que poderiam ser usadas na análise de dados. Estela mencionou que, se fosse a propor esta tarefa na aula, teria de a reformular. Na sua estruturação colocaria um conjunto de questões que indicassem explicitamente o que o aluno deveria fazer em cada uma delas, acrescentando: "No início uma pessoa tem de direcionar [as questões] para que… se eles responderem bem, nós podermos dizer 'muito bem!' e aquilo ouvir 'muito bem' dá vontade [aos alunos] de saberem mais (…)" (entrevista 1). Para Estela, no 10.º ano de Matemática A, é possível trabalhar-se o desenvolvimento da capacidade de interpretação de dados através das representações e também a capacidade de interpretar os conceitos estatísticos, à semelhança do que observou na ação de formação; no entanto, considera necessário que este assunto seja alvo de avaliação. Ou seja, que surja nos exames nacionais, tal como acontece com os assuntos lecionados no âmbito dos outros temas incluídos no programa de 10.º ano. Estela considera, pois, que a avaliação exerce influência nos processos de trabalho e no que o professor faz na aula: "pode-se melhorar a capacidade de interpretar na sala de aula… se os professores mudarem (…) se não deixarmos para o fim [o tema] e a Estatística tiver mais interesse em termos de avaliação" (entrevista1).

5.2.4 Discussão sumária dos dados à luz dos domínios do conhecimento didático

Conhecimento do ensino do tópico estatístico (CETE). A professora revela que trabalha a Estatística na aula à semelhança dos outros temas matemáticos, no sentido em que, na Matemática, há sobretudo preocupação de chegar-se à resposta certa, é uma ciência que lida com certezas. Apesar de entender a Estatística como uma área que também lida com incertezas, por exemplo, através da Estatística inferencial, considera que esta vertente não está presente no programa. Estela é da opinião que atualmente está mais à vontade a lecionar a Estatística comparativamente às primeiras vezes em que a lecionou, devido a ter vindo a aprofundar gradualmente os seus conhecimentos nesta temática, em particular, sempre que se prepara para a lecionar. As aulas de Estatística são desenvolvidas em torno da resolução de tarefas. Estela propõe tarefas que incluem um conjunto de dados (de tamanho moderado), questões objetivas, nas quais a interpretação de resultados e a sua discussão à volta do contexto considerado não é habitualmente privilegiada. Por exemplo, não costuma colocar questões que peçam a interpretação de um resultado. A calculadora gráfica é uma ferramenta útil no ensino do tema da Estatística para a obtenção dos cálculos e representações gráficas. As tarefas a que recorre são habitualmente selecionadas de manuais escolares que, na sua opinião, apresentam tarefas que incluem dados reais que permitem despertar mais a atenção dos alunos.

Apesar de possuir o conhecimento de que as investigações estatísticas não estão contempladas no programa escolar, a professora também tem a opinião que as investigações estatísticas ou atividades que possam envolver análises e comunicação dos resultados são atividades muito desafiantes para os professores que não têm experiência em lidar com elas. De facto, tais atividades exigem questionamento, discussão, interpretação e comunicação das conclusões, que são capacidades que não são frequentemente desenvolvidas nas aulas cuja temática é a Estatística.

Conhecimento do aluno (CA). Para Estela, de uma maneira geral, os alunos consideram o tema da Estatística acessível, pelo facto de a maioria dos assuntos lhes serem familiares e da Estatística ser habitualmente lecionada com foco na sua vertente processual, acabando por treinar bastante a construção de representações gráficas. Estela é também da opinião que esses conhecimentos adquiridos ao nível da construção de gráficos podem não ser mais utilizados pelos alunos na vida futura deles. Apesar de Estela ter pouca experiência com atividades investigativas nas aulas, considera-as difíceis para os alunos. Acha também que expressões tais como *procurar padrões*, *discutir o contexto da situação* e *analisar a variabilidade dos dados* na aula de Estatística não dizem nada aos alunos por não lhes serem minimamente familiares. Considera que os alunos têm dificuldades em lidar com muitos dados e têm uma certa tendência para realizarem análises de dados focadas nas representações gráficas, sem procurar, contudo, estabelecer comparações entre dados e sem comunicar os resultados.

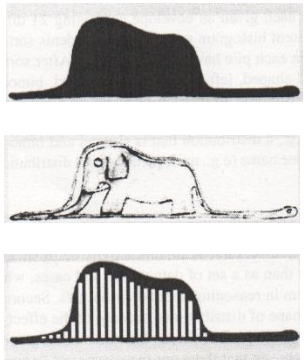
Conhecimento de Estatística (CE). Estela indica algumas limitações de alguns assuntos estatísticos do programa escolar e dá destaque ao papel da calculadora gráfica no ensino e aprendizagem do tema Estatística. A professora revela que gostaria de estar mais esclarecida relativamente ao processo de amostragem e à representatividade e dimensão de uma amostra.

Conhecimento do currículo (CC). Estela entende que o tema da Estatística do programa de Matemática A de 10.º ano se centra na Estatística descritiva, não incide na realização de inferências sobre uma amostra, nem na realização de investigações estatísticas; trata-se sobretudo de um tema que usa um conjunto de conceitos e representações, que são maioritariamente familiares aos alunos que iniciam o ensino secundário no 10.º ano. Deste tema curricular destaca como assuntos que os alunos irão lidar pela primeira vez no 10.º ano, o diagrama de extremos e quartis, as medidas de dispersão e o estudo das variáveis bidimensionais. Também aponta que a maioria destes assuntos podem ser expeditamente visualizados ou calculados com a ajuda da calculadora gráfica.

Para Estela, o ensino e a aprendizagem do tema da Estatística na Matemática A baseiam-se essencialmente numa faceta mais mecânica dos conceitos e na realização de representações gráficas. Na perspetiva da professora, o programa neste tema deveria ser mais desenvolvido, nomeadamente, incluir o aprofundamento dos conceitos e a realização de inferências, para que os alunos pudessem ver aplicações da Estatística a situações da vida real. Na sua opinião, esta seria uma forma mais desafiante para se dar continuidade ao tema da Estatística do programa do ensino básico (ME, 2007) (em vigor aquando da realização deste trabalho) que inclui vários conceitos e representações mencionadas no programa do ensino secundário (ME, 2001). Para se aprofundar mais o tema da Estatística no ensino secundário, o programa escolar deveria mudar nesse sentido e, simultaneamente, o tema deveria ser incorporado nos exames. Estela considera o tema Estatística da Matemática A relevante na educação escolar; contudo, o facto de ele ter pouca expressividade em termos de avaliação externa, nomeadamente nos exames do ensino secundário, condiciona o seu ensino, optando os professores, de uma maneira geral, por dedicar mais tempo a desenvolver todos os outros temas matemáticos que são objeto de avaliação nos exames nacionais.

5.3 A Estatística na sala de aula

A história do Principezinho de Saint-Exupéry começa com este desenho:
Para os adultos, este desenho era um chapéu.
Para a criança que o desenhou (o aviador, em criança) e para o Principezinho era uma jiboia que tinha comido um elefante.
Se o principezinho mostrasse o desenho original a um estatístico, ele diria: *Isto representa uma distribuição de dados*!



(Adaptado de Ben-Zvi & Garfield, 2008)

5.3.1 População e amostra

Já Cervantes (famoso escritor espanhol) afirmava
que de uma pequena porção se julga o todo:
não é preciso comer a sopa toda para saber se está salgada.
Um comentário às caraterísticas indutivas da Estatística, Pestana e Velosa (2010, p. 26)

Este tópico é analisado em três secções. A primeira (noções e dados estatísticos) incide sobre algumas noções centrais; é explicado como surgiram essas noções e é analisado como foram desenvolvidas na prática de Estela. Também inclui o tipo de amostras utilizadas no tópico e ao longo do tema. A segunda secção (reflexão sobre o tópico na aula) inclui algumas reflexões da professora sobre o ensino e aprendizagem do tópico e a terceira secção (discussão do tópico) apresenta uma discussão do tópico guiada pelas quatro dimensões do modelo do conhecimento didático do professor em Estatística.

5.3.1.1 Noções e dados estatísticos

Noções

Neste tópico foi proposta uma única tarefa na aula, a tarefa *Satisfação dos clientes*[16] (t1). Os conceitos e ideias intuitivas subjacentes foram sobretudo fornecidos pelos alunos que estiveram envolvidos na exploração e na apresentação desta tarefa à turma.

[16] Na aula, os alunos só tiveram acesso à tarefa que foi atribuída ao seu grupo. A maioria dos alunos tomou conhecimento da tarefa *Satisfação dos clientes* (t1) incorporada na *ficha desafio* durante a sua apresentação e discussão dinamizada pelo grupo responsável por ela. Todos os alunos tiveram um acesso posterior a essa ficha quando a professora a enviou para os respetivos e-mails no fim das apresentações.

Tarefa *Satisfação dos clientes* (t1)

A tarefa *Satisfação dos clientes* (t1) é uma das sete tarefas incluídas na primeira ficha de trabalho sobre Estatística, que a professora intitulou *Ficha desafio*, que foi distribuída para exploração e resolução a um grupo de quatro alunos. Esta tarefa tinha a particularidade de percorrer um conjunto de conceitos tais como população e amostra, unidade estatística, recenseamento e sondagem, representatividade de uma amostra e classificação da variável estatística em estudo (qualitativa/ quantitativa e quantitativa discreta/contínua).

Na aula, quando a professora pediu a este grupo para expor e discutir com a turma a sua resolução, uma aluna (Cátia) deste grupo[17] começou por apresentar a tarefa, lendo o enunciado, enquanto esta estava visível no retroprojetor; de seguida pediu a resposta à primeira alínea:

> **Cátia (grupo):** Alguém me sabe dizer qual é a população desta… desta tarefa?
> **Professora:** Deste estudo? (Corrige a professora)
> **Leonardo:** O cliente
> **Cátia (grupo):** A população... (é interrompida pela professora)
> **Professora:** Não digas [a resposta]! Só o Leonardo é que respondeu. Então, qual será a população?
> **Dalila:** Será… todos os clientes da empresa?
> **Professora:** Ah! (acenando com a cabeça)
> **Cátia (grupo):** (lendo do seu caderno) População é um conjunto de pessoas, objetos ou animais que tem uma caraterística em comum. E neste caso a caraterística comum é serem clientes da [empresa] *Preguicite aguda*. E alguém sabe o que é uma amostra?
> **Marco:** Amostra? Posso? [responder]
> **Cátia (grupo):** Podes.
> **Marco:** Do total de uma população… de um total de uma população recolhe-se uma… amostra? É isso?
> **Cátia (grupo):** (consulta a sua resposta no caderno e diz) A amostra é um subconjunto da população selecionada seguindo determinadas regras… para um específico objetivo, neste caso é para se saber a satisfação acerca do serviço prestado pela empresa. (observação 1)

Observa-se, neste excerto, que a professora tenta apoiar a aluna que está a apresentar a primeira questão de modo a conseguir que obtenha uma maior participação dos colegas da turma que assistem à apresentação. Os alunos Leonardo e Dalila indicaram corretamente a população do estudo. Marco, o único aluno que respondeu ao que era

[17] Nos excertos, optou-se por assinalar quais os elementos do grupo para os distinguir dos restantes alunos que dialogaram com eles.

uma amostra, revelou alguma dificuldade em defini-la, fornecendo uma resposta um pouco circundante. Contudo, na definição que deu, usou a palavra amostra no sentido de uma parte de um todo. A aluna Cátia do grupo referiu que a seleção dos elementos das amostras seguia "determinadas regras" mas não indicou nem foi questionada sobre quais seriam as possíveis regras na situação concreta. A aluna Cátia do grupo teve também o cuidado de ir transmitindo as definições dos conceitos que apareciam na tarefa. Esta atuação parece revelar que este grupo de trabalho também tinha a função de dar a conhecer as noções implicadas na tarefa.

A apresentação deste grupo de trabalho continuou da seguinte forma:

> **Cátia (grupo):** E alguém sabe-me dizer qual é a unidade estatística, sem ser os [alunos] que já participaram? **(**Ninguém responde)
> **Ana:** Eu não ouvi [a pergunta]…
> **Cátia (grupo):** Unidade estatística. Alguém me sabe dizer o que é uma unidade estatística, neste caso?... Eu vou dar uma definição de unidade estatística que pode ser que vos ajude… corresponde a cada elemento que constitui uma população. Logo quem é que sabe?... Quem é o génio?
> **Vários alunos: [**risos]
> **Leonardo**: O cliente!
> **Cátia (grupo):** Muito bem.
> **Cátia (grupo):** Depois temos que dizer se a amostra escolhida era ou não representativa do que estávamos a estudar. Alguém sabe o que é uma amostra representativa?
> **Marisa:** Se houve suficiente clientes [na amostra escolhida]… é se é suficiente para representar a população.
> **Laura (grupo):** É mesmo isso! Nós neste caso, a amostra representativa, nós não temos muito bem a certeza se é ou não porque não temos muitas informações acerca do número [total] de clientes que existem. Por isso nós acreditamos que seja… representativa, mas depende do universo em questão. Mas em princípio, o facto de terem sido escolhidos 100 clientes… já é um número considerável… e terem sido escolhidos aleatoriamente também ajuda… e então a gente dá a entender que a amostra é representativa. (observação 1)

A professora esteve atenta a esta interação mas não sentiu necessidade de participar nela. Neste excerto, a aluna Marisa, ao comentar sobre a representatividade da amostra em estudo apontou que a sua dimensão deveria ser "suficiente" para "representar a população". Na correção desta resposta, a aluna Laura (grupo) concordou prontamente com o que a Marisa disse e, ao argumentar sobre a possível representatividade da amostra dos 100 clientes da empresa, destacou a relevância de se conhecer o tamanho da população para se averiguar a adequação do tamanho da amostra. Embora, neste caso concreto, essa informação fosse desconhecida, a mesma aluna referiu que o seu grupo suspeitava que o tamanho da amostra fosse aceitável; o outro argumento usado por ela foi o facto de que, no processo seletivo desta amostra, ter sido usado amostragem aleatória, tal como o enunciado indicava. Todos estes aspetos foram referidos de um modo bastante superficial. Contudo, com esta tarefa parece ter ficado subentendido que a comparação do tamanho da amostra com o tamanho da respetiva população e, ainda, a sua aleatoriedade são condições a serem obrigatoriamente verificadas para uma amostra ser representativa. Estela não indicou que a primeira condição não era adequada.

A resolução desta tarefa prosseguiu com as intervenções da aluna Lurdes e do aluno Luís do grupo. A Lurdes tentou que lhe indicassem se a situação que se estava a tratar seria um censo ou uma sondagem e o aluno Luís pediu que lhe classificassem a variável em estudo, ou seja, se seria quantitativa ou qualitativa:

> **Lurdes (grupo):** Alguém sabe dizer se este estudo é um censo ou uma sondagem?
> **Vários alunos:** É uma sondagem…
> **Professora:** (Há barulho na sala com vários alunos a responder) Estou a ouvir grupos a responder em simultâneo…
> **Lurdes (grupo):** (que opta por fazer a pergunta a um aluno) Pedro, porque é que é uma sondagem?
> **Pedro:** Penso que o censo abarca tudo…
> **Lurdes (grupo):** O censo é mais dispendioso. Vou-vos explicar… por exemplo, o censo abrange toda a população… [através dele] pode retirar-se toda a informação sobre o indivíduo, enquanto que numa sondagem não… a sondagem é uma amostra da população… por exemplo, esta empresa *Preguicite aguda* não deve só ter 100 clientes… tem de ter mais clientes, não é?!… Fazer uma amostra desta população é uma sondagem.
> **Luís (grupo):** Alguém sabe-me dizer se a variável “satisfação” é uma variável quantitativa ou qualitativa?
> **Ana:** É quantitativa.
> **Luís (grupo):** Tens a certeza, Ana? (risos) É qualitativa! (de seguida, distinguiu variável quantitativa da qualitativa, referindo exemplos para cada situação). (observação 1)

Observa-se, mais uma vez, no excerto acima, que Estela deu uma orientação no sentido de ajudar o aluno do grupo a conseguir respostas dos colegas. As ideias que surgiram à volta de sondagem e de censo acabaram por ser mencionadas de uma forma superficial. A aluna Lurdes associa censo a um estudo que se baseia em toda a população enquanto a sondagem se baseia numa amostra. Deu também a entender que a amostra era uma sondagem, o que não era adequado. No entanto, a professora não interveio para fazer qualquer reparo e também não surgiram questões à volta destas noções.

De uma maneira geral, nesta tarefa, não houve muita participação da turma às questões solicitadas pelos elementos do grupo que notoriamente tinham a responsabilidade de dinamizar a discussão e de introduzir, de alguma forma, os conceitos envolvidos na tarefa. Só o grupo a quem foi atribuída a tarefa *Satisfação dos clientes* (t1) teve oportunidade de pesquisar sobre os assuntos nela incorporados, o que poderá ter condicionado a intervenção restrita que houve da turma nos momentos de discussão e correção da tarefa. É de destacar que esta foi a única tarefa que emergiu na sala de aula em que se abordou os conceitos de população, de amostra representativa e, ainda, os de censo e sondagem. Estes conceitos não voltaram a surgir na aula, ao longo do tema Estatística.

Recolha de dados e amostras na aula

A tarefa *Satisfação dos clientes* (t1) não incluiu a lista dos clientes da empresa, nem as percentagens dos clientes satisfeitos/não satisfeitos; apenas informou o tamanho da amostra aleatória considerada, correspondendo a 100 clientes. No entanto, a maioria das tarefas propostas nas aulas de Estatística incluiu conjuntos de dados que foram tomados como amostras e não fez referência à fonte dos dados nem ao método de recolha de dados.

Há três tarefas que foram elaboradas pela professora que explicitavam que os dados eram reais. Nomeadamente, na tarefa *Relatório sobre os resultados do teste intermédio* (t12), os dados escolhidos foram as notas (arredondadas às unidades) que os alunos da turma (de 10.º ano) tinham alcançado no teste intermédio recentemente realizado; na tarefa *A equipa de basquetebol do Porto* (t16), os conjuntos de dados foram retirados de um *site* desportivo, tal como indicado no seu enunciado; já na tarefa *Estudo com carateres qualitativos e quantitativos* (t20), os diferentes conjuntos de dados reais (qualitativos e quantitativos) foram formados com a recolha de informação junto da turma na aula. Mais especificamente, para as variáveis *peso* e *altura*, os alunos indicaram como valores as estimativas que fizeram, pois não havia nem balança nem fita métrica para se fazer as medições de forma precisa.

O entendimento de que há várias fontes de variabilidade nos dados não surgiu nas aulas de Estela. Em particular, não foi explorado que, na recolha de dados, essa variabilidade pode ser causada pela imprecisão do instrumento de medição ou por algum erro no registo de informação.

5.3.1.2 Reflexões sobre o tópico na aula

De seguida passa-se a expor e analisar as intenções da professora face à proposta da tarefa *Satisfação dos clientes* (t1), na aula, e também às perspetivas que a professora possui relativamente ao ensino do conceito de amostra e de amostra representativa.

Na segunda entrevista, referindo-se particularmente ao propósito da tarefa *Satisfação dos clientes* (t1), Estela disse tratar-se de uma tarefa que iria permitir, por um lado, a introdução de algum conteúdo estatístico que os alunos não tivessem aprendido no ensino básico (tal como: unidade estatística, amostra, amostra representativa) e, por outro lado, que os alunos recorressem ao manual à procura de definições e exemplos de diferentes tipos de amostras:

> Já tinha um bocadinho de [conteúdo] novo… fala em amostra, unidade estatística, amostra representativa. (…) [Nesta tarefa] Queria também que [os alunos] fossem ao livro [manual] e vissem alguns tipos de recolha de amostra. A estratificada, aleatória, aleatória estratificada, a que mais se usa é a estratificada… sistemática, por grupos… há várias. (entrevista 2)

De acordo com a professora é na disciplina de MACS[18] que a amostra é estudada de modo mais profundo, em comparação com a disciplina de Matemática A, abrangendo mais conceitos e aplicações. Estela indicou, a título de exemplo, que na Estatística inferencial que ensina em MACS, realiza tarefas cujo objetivo é identificar e distinguir entre parâmetros e estatísticas (algo que não propõe na disciplina de Matemática A); esta distinção tem obrigatoriamente em conta se se está a lidar com a uma população ou com uma amostra.

Estela também referiu que foi através de MACS que começou a melhorar os seus conhecimentos estatísticos relativamente a amostras e afins, revelando algum descontentamento relativamente ao grau de profundidade com que a noção de amostra aparece especificada no programa de Matemática A no 10.º ano. Na sua opinião, o

[18] O programa de MACS (ME, 2001) não faz referência a métodos que permitam determinar a dimensão de uma amostra, embora dê algumas indicações sobre o tamanho da amostra para a determinação da distribuição de amostragem de uma estatística.

programa deveria abranger a questão da dimensão da amostra, da sua representatividade e outras noções associadas a ela, tais como, erro, nível de confiança e intervalo de confiança:

> Mas por exemplo, tenho muita pena que no [programa] 10º ano Matemática A, não se fale na dimensão da amostra, qual deve ser o tamanho da amostra para que ela seja representativa?... Não se fala em intervalo de confiança, nível de confiança, erro amostral. Isso foi uma coisa que aprendi com Matemática Aplicada às Ciências na preparação e durante a sua lecionação. (entrevista 1)

Quando foi questionada sobre possíveis ideias intuitivas que os alunos (de Matemática A) retinham, de uma maneira geral, sobre amostra representativa, Estela confessou:

> Eles não têm ideia do que é uma amostra representativa. Nem eu tinha como professora, nem quando fiz o mestrado. Só depois é que fui lidando com isso aos poucos… E com [a disciplina de] Matemática Aplicada às Ciências Sociais é que eu fui aprofundando (…) Mas a ideia de que um 10.º ano de Matemática A tem é a ideia de proporção. (entrevista 3)

Para Estela o conceito de amostra representativa foi-se desenvolvendo ao longo da sua prática profissional. Baseando-se na sua experiência profissional, acha que os alunos, quando pensam em amostra representativa, ponderam sobretudo o aspeto dimensional da amostra, estabelecendo uma certa proporção que lhes pareça razoável entre o tamanho da população e o tamanho da amostra em causa. Para a professora esta ideia está, de certo modo, em sintonia com o programa escolar.

Ao ser mais explícita acerca do que os alunos pensam sobre este conceito, Estela considera que, para os alunos, se o tamanho de uma amostra for cerca de 1/3 ou 1/2 do tamanho da população esta estará com um tamanho aceitável para ser representativa:

> Por exemplo, numa turma de 25 alunos, se fizerem questões a dez [alunos]… acho que já devem achar que é representativa, ou seja, a ideia deles de dimensão deve ser de 1/3 da população que tiverem ou até 1/2 (…) Eles têm noção se em 25 alunos se fizerem um questionário a um, esta amostra não é representativa da turma mas se já fizerem a dez ou quinze já é muito bom porque é quase a turma toda. (entrevista 3)

Notoriamente, para a professora e alunos, uma amostra maior terá sempre maior probabilidade de ser representativa do que uma amostra mais pequena. Estela não tem dúvidas que os alunos, de uma maneira geral, sabem que as amostras representativas têm de ser aleatórias, acabando por exprimir também o conceito que ela tem de amostra aleatória:

> Têm a noção de que a amostra [representativa] deve ser aleatória, todos [os elementos da amostra] tem a mesma probabilidade de serem selecionados… [amostra] não é intencional… ou por conveniência, é na aleatória que eles pensam. (entrevista 3)

Ao comentar sobre o facto dos alunos do grupo terem tido uma participação bastante organizada na discussão da tarefa e ter havido o cuidado de cada aluno do grupo participar na apresentação de alguma parte da tarefa, Estela mencionou que não foi a primeira vez que tinham realizado trabalhos de grupo. Também informou que teve algum cuidado em impor e esclarecer certas normas de conduta em grupos (repartição de tarefas no grupo, comunicação ordeira dentro do grupo, realização dos registos das respostas finais, apresentação dos trabalhos aos colegas, etc.) desde o início do ano letivo, nomeadamente, no tema da Geometria.

Nas diferentes entrevistas, Estela revelou o interesse pessoal em desenvolver mais o seu conhecimento estatístico, nomeadamente, sobre amostragem e dimensão da amostra. Confessou que gostaria de poder levar à sala de aula, pelo menos, uma situação, que pudesse ser trabalhada junto dos alunos do secundário, que envolvesse uma recolha de dados adequada com tomada decisões sobre a amostragem a realizar e o tamanho da amostra. Contudo, acha que o programa de Matemática A não abrange o desenvolvimento deste tipo de atividades, que classificou como sendo da área da Estatística indutiva.

5.3.1.3 Discussão do tópico

CETE. Estela propôs apenas uma tarefa que abordava o tópico População e amostra, cujo enunciado incluía informação minimal acerca dos dados e seu contexto. Além disso, a sua exploração ficou incumbida apenas a um grupo de quatro alunos da turma. Os restantes alunos só tiveram oportunidade de lidar com alguns conceitos novos e com a tarefa em si no momento em que grupo responsável por ela a apresentou. Houve definições, como a de amostra representativa ou a de amostra aleatória, que acabaram por não ser explicitamente dadas. Mesmo a professora não sentiu necessidade de intervir para questionar os assuntos estatísticos trabalhados na tarefa durante a interação dinamizada pelo grupo de alunos ou para esclarecer alguma situação referida.

Embora tenha usado algumas amostras com dados reais nalgumas tarefas do tema Estatística, não fez notar a relevância de se controlar nelas a *variabilidade*. Todas essas situações apontadas mostram indícios da leitura que Estela faz do programa escolar no tópico.

CA. Na aula de Estela houve, de uma maneira geral, um acesso limitado aos raciocínios dos alunos da turma em torno da única tarefa trabalhada. Contudo, a professora achou que os alunos do grupo tiveram uma boa prestação na apresentação e na sua correção. Durante essas apresentações tentou, sobretudo, ajudá-los para que conseguissem uma maior participação dos colegas às questões que lhes colocaram. Na correção desta tarefa destaca-se o facto da aluna Laura do grupo afirmar que não tinham a certeza da amostra referida com 100 clientes ser representativa, por não haver indicações sobre o tamanho da população, mas que o seu grupo supunha que o fosse. No entanto, a professora não fez nenhum comentário relativamente a esta situação. A ideia principal que foi comunicada pelo grupo acerca de amostra é a de que uma amostra deve ser representativa e aleatória e de que a sua representatividade deve ser avaliada tendo conta uma certa proporção entre o tamanho da amostra e o da respetiva população. Esta ideia está em consonância com o conhecimento do aluno sobre este assunto que a professora revelou na entrevista final. Ou seja, de acordo com a sua experiência, os alunos neste nível de ensino possuem a ideia de que uma amostra representativa é aleatória e o seu tamanho deve ser de pelo menos 33% da população considerada.

CE. Estela tem a convicção de que as suas experiências na lecionação da Estatística na disciplina de MACS foram essenciais para o desenvolvimento dos seus conhecimentos sobre este tópico e, de um modo mais abrangente, sobre Estatística inferencial. Revelou, embora de uma forma superficial, conhecer alguns tipos de amostragem (estratificada, sistemática, de conveniência) e ter noção de que obter-se uma amostra representativa é um assunto que se reveste de alguma complexidade. Na sua opinião, amostras aleatórias maiores têm maior probabilidade de ser representativas do que as mais pequenas. No entanto, parece descurar a influência da variabilidade dos dados na escolha da dimensão da amostra. Assim, a professora considera válido que se julgue simplesmente a representatividade de uma amostra através da sua dimensão, correspondendo a uma determinada percentagem da população de valor igual ou superior a 33%.

CC. Para Estela, os assuntos do programa de Matemática A relacionados com população e amostra devem ser mencionados de um modo bastante superficial, uma vez que considera que este programa não vai além da Estatística descritiva, não incluindo a realização de inferências estatísticas. Este conhecimento curricular de Estela tem inevitavelmente repercussões ao nível do seu CETE.

5.3.2 Organização e interpretação de dados

Readers bring their background knowledge and experiences
to bear when processing information,
whether they are reading prose, tables or graphs.
Curcio e Artz (1996, p. 670)

A análise que se faz do conhecimento didático em Estatística de Estela, relativa a este tópico, tal como refletido na sua prática, divide-se em três secções. Na primeira (representações) apresenta-se como é que as tabelas de frequências foram introduzidas e trabalhadas nas aulas e mostra-se como foi desenvolvida a função cumulativa, em particular, as relações estabelecidas desta função com outras representações. Na segunda secção (reflexão sobre o tópico na aula) inclui-se reflexões da professora sobre o ensino e aprendizagem do tópico. Na terceira secção (discussão do tópico) realiza-se uma discussão em torno da *Organização e interpretação de dados*, guiada pelas quatro dimensões do modelo do conhecimento didático do professor em Estatística.

5.3.2.1 Representações

Estela propôs algumas tarefas na sala de aula que envolviam conjuntos de dados numéricos ou dados representados numa representação gráfica (por exemplo, gráfico circular, gráfico de barras ou diagrama de caule e folhas) que teriam de ser representados numa tabela com frequências absolutas e relativas e com as respetivas frequências acumuladas. Os enunciados das tarefas abrangidas neste tópico (em particular, *Número de cães examinados* (t3), *Ração gasta num dia* (t6) e *Classificações a Matemática das turmas A e B* (t4) da *ficha desafio*) não solicitavam o recurso à calculadora gráfica. Contudo, Estela pediu à maioria dos grupos de trabalho, constituídos na aula, para explorarem e resolverem as tarefas propostas na *ficha desafio*, com o apoio da calculadora gráfica. Nas aulas acabaram por surgir representações diversificadas para os mesmos conjuntos de dados, à medida que os alunos foram descobrindo como as fazer e exibir nas suas calculadoras gráficas[19], como se irá mostrar.

Tabelas de frequências

Tarefa *Número de cães examinados* (t3)

Depois de apresentar a tarefa à turma, o aluno Mário do grupo mostra a tabela de frequências (figura 14) totalmente preenchida.

Nº cães	Freq. Absolutas	Freq. Relativas (%)	Freq. Absolutas Acumuladas	Freq. Relativas Acumuladas (%)
15	4	16	4	16
20	5	20	4+5=9	16+20=36
25	6	24	9+6=15	36+24=60
30	4	16	15+4=19	60+16=76
35	3	12	19+3=22	76+12=88
40	3	12	22+3=25	88+12=100

Figura 14: Tabela de frequências da tarefa *Número de cães examinados.*

Quanto ao preenchimento da coluna das frequências absolutas acumuladas na tabela, houve a seguinte interação:

> **Mário (grupo):** (acerca das frequências absolutas acumuladas diz) Na frequência acumulada é somar todos os parâmetros de frequências absolutas, a soma tem de dar 25… os 25 dias.
> **Lurdes:** Como é que somaste?

[19] As instruções comunicadas na aula diziam, na maioria das vezes, respeito à da calculadora *Texas Instruments*, pelo facto dos alunos estarem quase todos equipados com a calculadora gráfica dessa marca.

> **Mário (grupo):** Somei, inicialmente, 4 mais 5… o 4 é o número de dias que aparece 15 e o 5 é o número de dias com 20 [cães examinados], quatro mais cinco dá 9. As somas anteriores vão para baixo e adicionamos o que há nessa linha (ao explicar o 9+6=15 da 3.ª linha da coluna das frequências absolutas acumuladas)… (observação 1)

O aluno Mário do grupo tenta explicar que somas é que têm de ser efetuadas para se completar as frequências absolutas acumuladas integrando o que entende, no contexto da situação, por cada valor utilizado. A professora não interveio nesta parte.

Após esta interação, o aluno Marco do grupo mostrou à turma como chegar às frequências absolutas e às frequências absolutas acumuladas e suas respetivas frequências relativas, através das listas da calculadora gráfica. Para as frequências acumuladas (absolutas e relativas) usou a instrução *2nd LIST*, selecionando *cumSum* em *OPS*. Este aluno também mencionou que, no seu grupo de trabalho, não tinham conseguido descobrir como fazer o gráfico de barras na calculadora gráfica. A professora pediu-lhes que fizessem uma nova tentativa, agora com a calculadora projetada no quadro, enquanto faziam a apresentação, informando que os ajudaria caso não conseguissem chegar à representação. Um dos primeiros gráficos obtidos foi o da figura 15. A professora alertou que este gráfico não correspondia à distribuição do número de cães tratados por dia pois “as frequências não são todas iguais” (observação 1), e pediu-lhes para confirmarem as listas selecionadas para o gráfico. Estela detetou, com prontidão, que a dificuldade deles estava na inserção das listas corretas e no ajuste da janela.

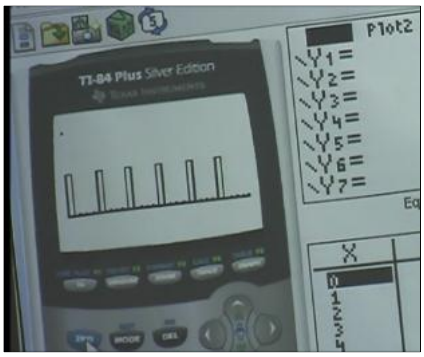


Figura 15: Gráfico de barras inicial.

O gráfico correto (figura 16) foi obtido depois da seleção correta das listas e do ajuste feito à janela de modo a esta englobar todos os valores assumidos pela variável em estudo. Estela também solicitou uma leitura ao gráfico de barras da distribuição do número de cães tratados (ver, também, excerto seguinte):

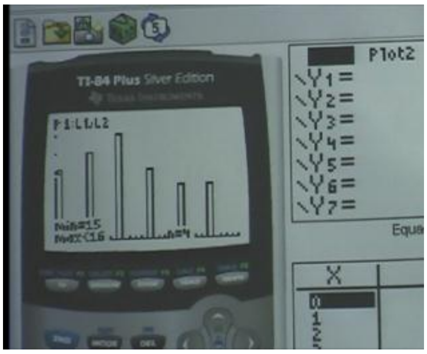


Figura 16: Gráfico de barras.

> **Professora:** Têm mais alguma questão? (ao que nenhum aluno responde) Então tenho eu! (faz um sorriso)… Façam o *TRACE* sobre o gráfico… (E, referindo-se à primeira barra observada na calculadora gráfica)… O valor da unidade estatística é 15… aí houve 4 dias (n=4) que foram observados quantos cães? O *TRACE* ao estar sobre a primeira barra mostra as informações, *n*=4 dias e *min*=15 número de cães tratados. Então, quantos cães foram observados?
> **Mário (grupo):** É 16.
> **Professora:** 16? Não, foi 15!! Façam a seta ir para a direita. Estão a ver?… Ela assim vai para a outra barra e dá também as informações… (observação 1)

Para se ter um acesso mais imediato à escala do gráfico que não surgia no ecrã, Estela pediu aos alunos para clicarem na tecla *TRACE,* que fazia aparecer no ecrã o valor da variável através de *min=valor numérico* e a sua respetiva frequência através de *n=valor numérico* sobre cada barra do gráfico de barras. Ao ter detetado que o aluno respondente, Mário (do grupo), não estava a fazer corresponder o parâmetro correto à variável em estudo, o significado de cada parâmetro (*n*, *min* e *máx*) foi algumas vezes explicado pela professora, à medida que percorreu algumas barras do gráfico exposto.

Tarefa *Ração gasta num dia* (t6)

Esta tarefa também pertencia à *ficha desafio*. Na sua apresentação, há o aluno Hugo do grupo que exibe a tabela de frequências pedida no enunciado, totalmente preenchida (figura 17), com um breve comentário, no qual esclarece em que consiste cada coluna.

Estela reagiu ao facto dos alunos do grupo optarem por realizar a contagem do número de dados inferiores a 32,1 quilogramas através da lista original de dados (alínea d) em vez de tirarem partido da tabela de frequências que já tinha sido completada por eles (alínea anterior, c):

Ana (grupo): Temos de contar, agora, os valores que foram inferiores a 32,1.
Dalila: É o 30,3; 25,8; 20,3; 21,9; 19,2; 27,3; 19,7; 29,2
Jaime: São 8!
Marisa (grupo): Exatamente, são 8 valores!
Professora: Ora bem. Posso dizer uma coisa? Depois de fazerem a tabela…. Por alguma razão as questões e as alíneas são seguidas… Acham que tínhamos necessidade de andar a contar os dados todos outra vez na lista acima? Não! Porque eles estão aqui… (apontando na tabela para o valor da terceira linha das frequências absolutas acumuladas visto corresponder à classe com limite superior de valor 32,1 quilos)
Dalila: Ah, pois é!! (ouve-se vários risos)
Professora: E agora, para dar significado à coluna das frequências absolutas acumuladas. Qual a informação que ela nos dá? ... A primeira linha da coluna (absolutas simples) diz-nos a mesma coisa que esta linha (absolutas acumuladas), ou seja, diz que existem três dias em que foram gastos … a quantidade de ração que foi gasta está neste [primeiro] intervalo, certo? Mas esta linha [segunda] já diz o quê?
Ivo: o total.
Professora: O total… Diz que em cinco dias foi gasta ração, de um aviário, entre quantos?
Professora e alunos: [em simultâneo] 19,5 e 23,7.
Professora: E em oito dias?
Professora e alunos: entre 19,5 e 32,1.
Professora: Portanto, [a tabela] dá-nos logo essa informação direta sem estar a preocupar e ir contar outra vez [nos dados iniciais]. Pronto!
(observação 1)

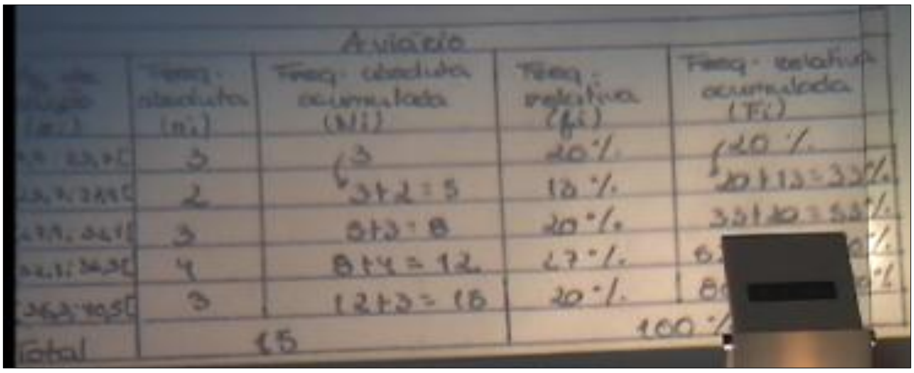


Figura 17: Tabela de frequências.

Nesta interação, Estela deu sentido aos números das frequências absolutas acumuladas, interligando-os com o contexto. Também explicou que em oito dias se gastou menos de 32,1 quilogramas de ração num aviário através da tabela de frequências absolutas acumuladas, como alternativa à estratégia usada pelo grupo.

Tarefa *Classificações a Matemática das turmas A e B* (t4)

Esta tarefa estava também incluída na *ficha desafio*. Na sua apresentação e discussão inicial pelo grupo que a estudou não foi usada a calculadora gráfica, certamente porque a tarefa pedia um diagrama (de caule e folhas) que não é feito nas calculadoras gráficas e, além disso, incorporava duas questões que implicavam contagens e cálculos básicos. No decorrer da mesma aula, dois alunos deste grupo (Leonardo e Dalila) ofereceram-se para exibir cálculos e gráficos que descobriram, ao longo da aula, na calculadora gráfica *TI-Nspire* com os dados da turma A desta tarefa.

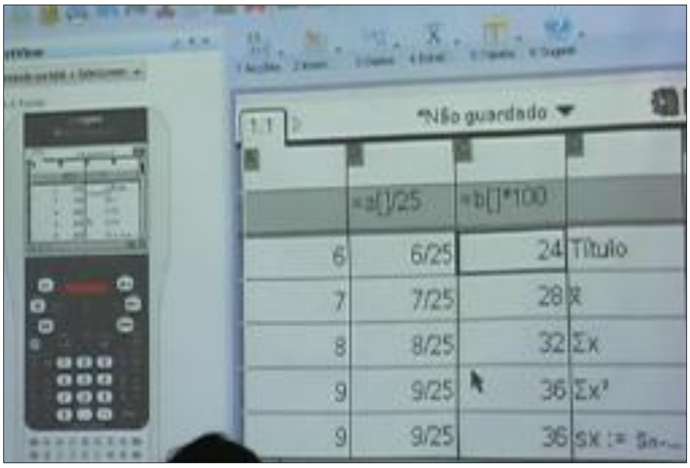


Figura 18: Folha de cálculo na *TI-Nspire*.

Dalila introduziu as 25 notas na coluna A, da folha de cálculo desta calculadora gráfica (*TI-Nspire*) projetada (ver figura 18). Ao realizar esse processo, colocou, por lapso, a nota 155 em vez da nota 15. Há uma aluna que alerta Dalila para este erro e que também lhe sugere que poderia colocar os dados em apenas duas colunas (uma com os valores da variável e outra com os valores das respetivas frequências absolutas). A professora achou a sugestão passível, mas também referiu que, para isso, era preciso despender tempo a sintetizar os dados à mão, numa tabela. Dalila também mostrou como determinar de uma só vez "estatísticas de uma variável" (média, mediana, quartis) do *menu estatística* (que colocou na coluna D) depois de assinalar qual a coluna dos dados e de indicar que a frequência de cada dado era 1. Adicionalmente, esta aluna colocou na coluna B as frequências relativas de cada nota, dividindo cada frequência por 25, e na coluna C as frequências relativas mas em percentagem dizendo aos colegas que "podem fazer a equação que traduz a frequência relativa, a/25, usando esta fórmula na folha de cálculo" (observação 2). Nem a professora nem os alunos se aperceberam que essas frequências relativas estavam erradas visto que as notas não estavam agrupadas pelas suas frequências absolutas; assim, o valor cada uma delas deveria ser 0.04 ou 4 (%) em vez dos diversos valores obtidos na coluna C.

Por sua vez, o Leonardo, com os mesmos dados da turma A, fez e mostrou as cinco representações abaixo (figuras 19 e 20). Destacou, contudo, que não entendia o diagrama de pontos (figura 19) e que só conseguiu fazer o diagrama circular e o de barras (figura 20) porque usou a opção *forçar categórico* do *Menu dados e estatística*, em *propriedades dos gráficos* (observação 2). A professora explicou que o diagrama de pontos "mostra cada uma das notas" e, quanto aos outros dois gráficos, pareceu-lhe que a calculadora gráfica tinha a opção de transformar dados numéricos em dados qualitativos, sendo que cada "nota diferente corresponde a uma categoria" cujas densidades surgiam representadas no gráfico de barras por uma barra e no gráfico circular por um setor (observação 2).

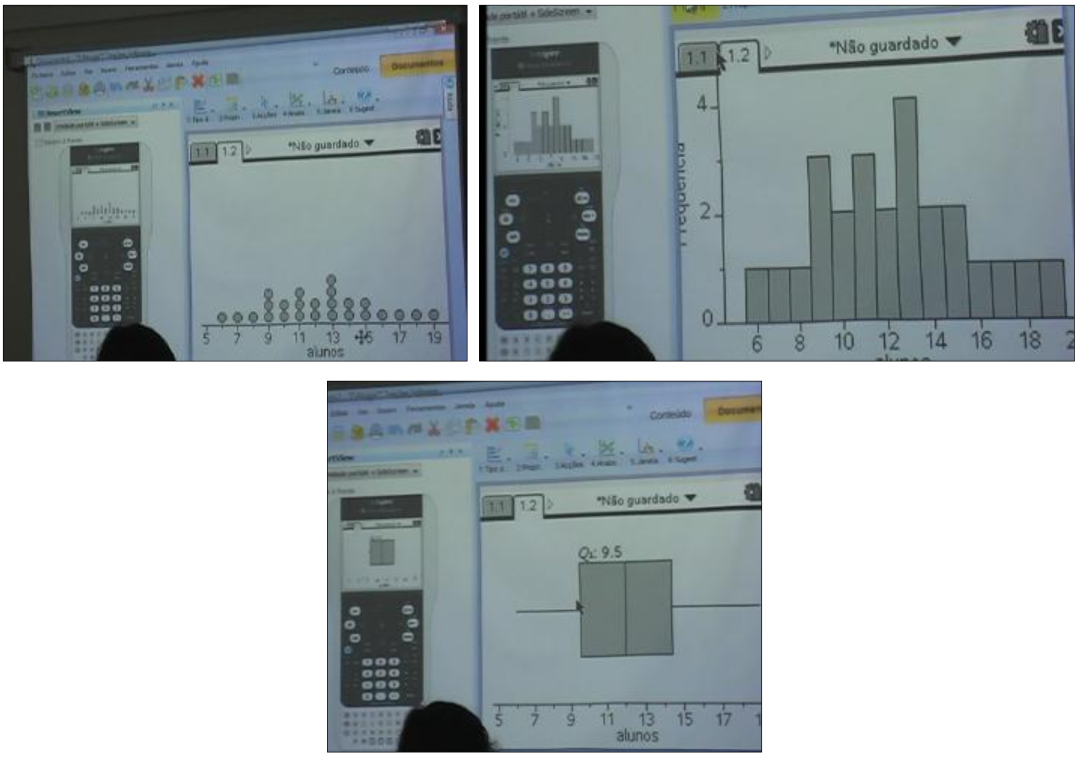


Figura 19: Gráfico de pontos e histograma (acima) e diagrama de extremos e quartis (abaixo).

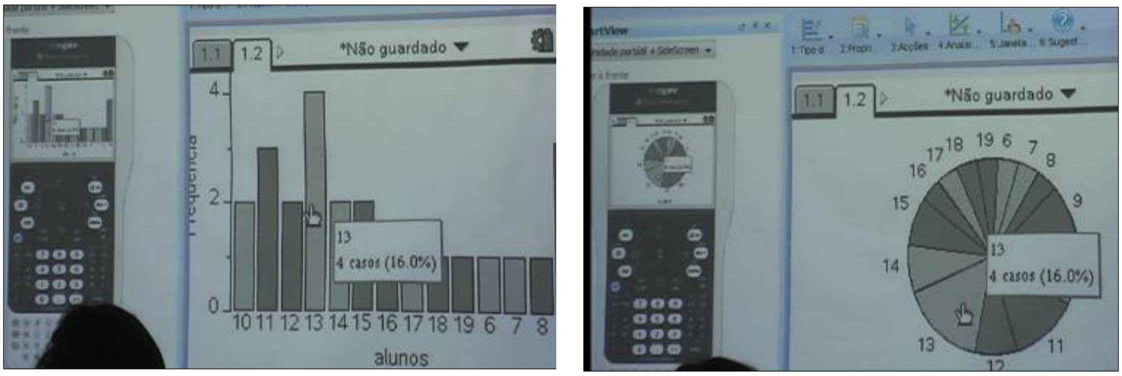


Figura 20: Um gráfico de barras e o diagrama circular.

Durante esta mostra de representações pelos dois alunos não houve mais questões colocadas pela turma nem pela professora, além das colocadas pelos alunos que exibiram as representações. Foi possível observar cinco representações do mesmo conjunto de dados, algumas delas a reiterar caraterísticas comuns e outras a evidenciar caraterísticas distintas dos dados, mas na aula não houve discussão destes aspetos. Foi também a primeira vez que, na aula, surgiu um gráfico de pontos e que os alunos viram representado numa calculadora gráfica o gráfico circular (ver figuras 19 e 20).

Função cumulativa

Para abordar a função cumulativa, a professora apoiou-se numa situação concreta do manual que incluía a expressão analítica de uma função cumulativa e a sua representação gráfica. A tarefa *Idades de jovens num casting* (t5) foi a única tarefa proposta na aula que versava especificamente a função cumulativa.

Tarefa *Idades de jovens num casting* (t5)

Na resolução desta tarefa, Estela pediu aos alunos para consultarem a página do manual com a situação introdutória que, segundo lhes explicou, retratava uma função cumulativa, para, a partir dela, tentarem reproduzir a função cumulativa pedida no enunciado da tarefa t5. Nesta tarefa, os dados foram inicialmente representados num diagrama de caule e folhas e, a partir desse diagrama, a professora fez a tabela no quadro com a ajuda dos alunos (registando frequências absolutas e frequências absolutas acumuladas). Estela descreveu a função cumulativa como uma função "que representa as frequências absolutas acumuladas" (observação 2). Indicou, de igual modo, que era possível determinar a sua expressão e proceder à sua representação gráfica.

Depois de registar no quadro a expressão analítica da função cumulativa relativamente aos dados desta tarefa, pediu a um aluno para fazer a sua representação gráfica no quadro (ver figura 21). Adicionalmente, Estela questionou os alunos sobre o significado de *F(x)*, concretizando alguns valores de *x*:

> **Professora:** O que é que significa por exemplo *F(21)*? A variável estatística é a idade dos jovens…
> **Ivo:** *F(21)* é o número de participantes com 21 anos.
> **Professora:** Mas, foram assim tantos?
> **Ivo:** Não… [é o número de participantes] com idade inferior ou igual…
> **Professora:** Ah!... Com idades iguais ou inferiores a 21… Não há uma maneira mais bonita de dizer isso?
> **Alguns alunos e professora:** [participantes] com idades inferiores a 22 anos!
> **Professora:** E quantos participantes tinham menos de 21 [*F(20)*]?
> **Jaime:** 18…

Professora: Estão a ver [com *F(x)*] dá para responder rapidamente às questões que envolvem contar até lá, ou seja, até ao valor de *x* considerado… Também dá para… reparem… será que dá para descobrir o aumento do número de pessoas de um ramo para o outro?.... Já ouvi [alguém dizer]!!... Só com uma…
Cátia: Não dá…
Professora e Lara: Subtração…
Professora: Por exemplo… 25-18 dá 7… que são os de 21 anos [*F(21)-F(20)=7* corresponde ao número exato de participantes com 21 anos] Pronto! (observação 2)

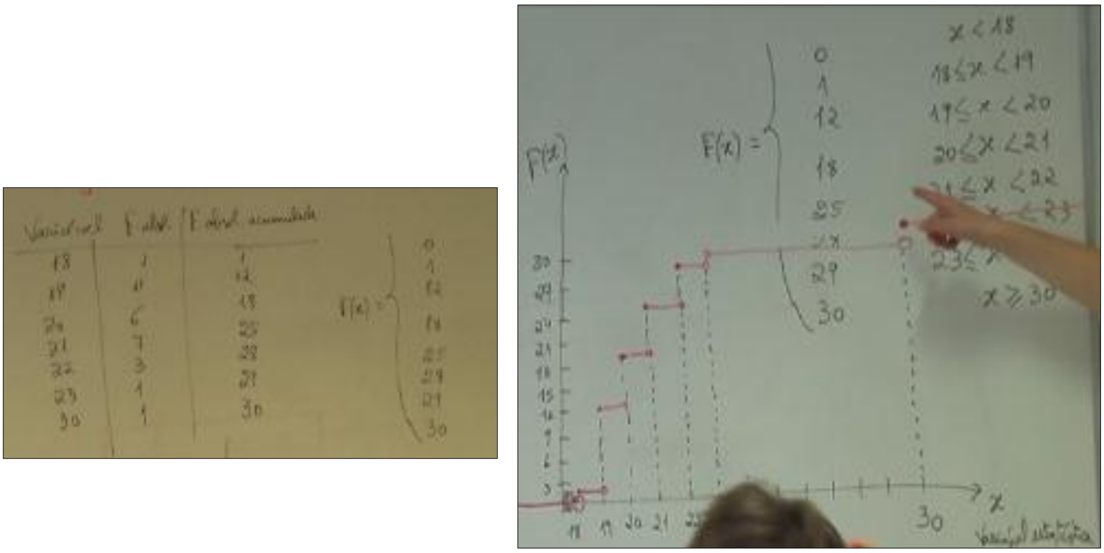


Figura 21: Tabela de frequências, função cumulativa *F(x)* e gráfico de *F(x)* no quadro.

Nesta interação, a professora tenta que os alunos percebam o significado de *F(x)* para alguns valores de *x* tendo em conta o contexto dos dados. Há alguns alunos que indiciam estar a acompanhar as suas explicações. Estela também explanou que, na expressão analítica da função cumulativa, de ramo a ramo, havia acumulação de frequências e que era possível obter as frequências não acumuladas (absolutas) quando se fazia subtrações entre os valores de ramos consecutivos.

5.3.2.2 Reflexões sobre o tópico na aula

Estela considera que representar dados em tabelas é algo que os alunos têm vindo a experienciar desde o ensino básico, e realça que eles estão bastante familiarizados sobretudo com as frequências absolutas. Na sua opinião, neste tópico, do ensino secundário, insiste-se na construção das tabelas de frequências relativas e relativas acumuladas, com o recurso imprescindível da calculadora gráfica, e na interpretação dos seus valores (entrevista inicial). Na entrevista inicial, Estela afirmou que não via "grande aplicabilidade da função cumulativa no caso dos dados quantitativos discretos, mas que, no caso contínuo, poderia ajudar a identificar os valores dos quartis". Uma vez que a função cumulativa é referenciada no programa escolar, Estela optou por trabalhar uma situação para que os alunos tivessem algum contacto "com a expressão e sua representação gráfica, que está associada às frequências absolutas e acumuladas" (entrevista 2).

5.3.2.3 Discussão do tópico

CETE. No ensino deste tópico foi dada primazia a que os alunos ficassem a entender como determinar frequências (absolutas/relativas) acumuladas e a sua utilidade para se responder a algumas questões sobre os dados, e a conhecer as instruções da calculadora gráfica para determinar frequências e representações gráficas. Estela não deixou de estabelecer a ponte entre frequências absolutas acumuladas e a função cumulativa e de dar oportunidade aos alunos para mostrarem as capacidades da calculadora gráfica *TI-Nspire* na representação de gráficos. Esta última situação permitiu, por um lado, que surgissem representações novas mas, por outro lado, não se relacionou mais aprofundadamente os gráficos, analisando-se o que cada um deles permitiria inferir sobre os dados, com base nas medidas de localização, que tinham sido referidas na aula. Apesar de a professora ter referido, na entrevista inicial, a relevância da função cumulativa na localização dos quartis para dados agrupados, na aula não propôs nenhuma tarefa ou situação que abarcasse a função cumulativa com esse tipo de dados.

CA. Estela foi detetando algumas dificuldades (por vezes detetadas pelos próprios) nos alunos em usar a calculadora gráfica, e também em utilizar a informação contida numa tabela de frequências para responder às questões colocadas sobre os dados. Estela deu também algum *feedback* às dúvidas que dois alunos lançaram sobre as representações que eles mostraram na *TI-Nspire*, para o mesmo conjunto de dados. Não detetou o erro que estes alunos fizeram quando pretendiam mostrar como fazer as frequências relativas na calculadora. Possivelmente se esta coluna errada tivesse sido usada mais vezes, por exemplo como componente de alguma representação, talvez o erro tivesse sido identificado. De um modo geral, registou-se um acesso limitado ao raciocínio dos alunos no desenvolvimento da maioria das tarefas visto estas terem sido exploradas e trabalhadas essencialmente por um grupo de alunos e com foco nos processos (cálculo e representações tabulares e gráficas) em detrimento da realização de interpretações e reflexões acerca do conjunto de dados.

CE. Este tópico, por ser sido trabalhado, na aula, de uma forma expedita, não permitiu capturar muitos aspetos desta dimensão do conhecimento didático da professora. Na aula, foi visível o seu à vontade em responder a dúvidas relacionadas com a utilização das calculadoras gráficas e em fornecer algumas explicações relativamente às representações que nelas eram obtidas, mesmo tendo em conta que não tinha grande experiência com a *TI-Nspire*.

CC. Para este tópico, Estela tinha claramente o objetivo de trabalhar alguns conceitos e representações com o apoio da calculadora gráfica. Foi relembrada a construção de tabelas, mostrado como fazer a leitura das frequências acumuladas, mesmo através da expressão analítica da função cumulativa e do seu gráfico. O trabalho desenvolvido na aula expressa a leitura de Estela sobre o programa relativamente a este tópico.

5.3.3 Medidas de localização

A statistician sees group features such as the mean and median as indicators of stable properties of a variable system – properties that become evident only in the aggregate.
Konold e Pollatsek (2004, p. 172)

Este tópico é analisado em três secções. Na primeira secção (média, moda e quartis) começa-se por fazer referência às situações em que as medidas de localização surgem em simultâneo na aula e é também analisado como é que essas medidas foram desenvolvidas na prática da Estela. Na segunda secção (reflexões sobre o tópico na aula) apresenta-se algumas reflexões da professora sobre o ensino e aprendizagem do tópico. Na terceira (discussão do tópico), e última secção, mostra-se uma discussão do tópico guiada pelas quatro dimensões do modelo do conhecimento didático do professor em Estatística.

5.3.3.1 Média, moda e quartis

Estela apresentou algumas tarefas que requeriam a cálculo das três medidas de localização média, moda e mediana em simultâneo. Os valores da moda, média e dos quartis foram, na maioria das vezes, determinados através da calculadora gráfica depois de serem inseridas as listas dos dados numéricos e de seguirem as instruções: em *STAT* (do *menu CALC),* selecionar *1-Var-Stats,* adicionar a uma lista de dados ou duas listas (caso houvesse informação sobre as frequências absolutas) e, depois procurar o valor da moda, da média aritmética ou da mediana no conjunto de medidas estatísticas obtidas na calculadora com esses passos. De modo a que os alunos automatizassem e memorizassem o procedimento, Estela teve o cuidado de repetir por várias vezes essas instruções, quer oralmente, quer ilustrando visualmente os passos instrucionais referidos acima, com a calculadora gráfica projetada na tela. Apenas nas primeiras tarefas é que esses cálculos e a identificação de estatísticas foram também efetuados à mão com conjuntos de dados relativamente pequenos. A atividade de estimar ou identificar o valor destas medidas de localização em representações gráficas, não emergiu na sala de aula com a mesma intensidade do que o cálculo destas medidas.

Média

Nesta secção mostra-se como é que a média surgiu pela primeira vez, na aula, no tratamento de um conjunto de dados quantitativos discretos (tarefa *Número de hóspedes de uma estância balnear*, t7) e também quando se esteve perante um conjunto de dados quantitativos contínuos (tarefa *Ração gasta num dia,* t6). Exibe-se ainda como é que valores de média foram associados a distribuições de dados representados em diagramas de barras (Tarefa *Distribuições das idades das equipas de trabalho*, t10).

A média é um conceito que, na sala de aula de Estela, não parece precisar de grandes apresentações, na medida em que surge associada ao algoritmo (média aritmética) e parece ser a mais familiar das três medidas de localização (média, moda e mediana) para os alunos. A leitura do seu resultado foi realizada, na maioria das vezes, sem se tomar em consideração o contexto dos dados. É também, das medidas de localização, a mais solicitada nas tarefas propostas na aula.

A explicação dada pelo aluno Jaime à turma, quando referiu o procedimento que o seu grupo efetuou para chegar ao resultado da média (alínea a) na Tarefa *Número de hóspedes de uma estância balnear* (t7, *ficha desafio*), foi a seguinte:

> Nós colocámos os números que aparecem aqui (apontando para o acetato com a resposta escrita, referindo-se à soma dos 7 dados numéricos discretos) e dividimos pelos dias da semana (aponta para o número 7, denominador da fração)… dá o valor aproximado por dia, ou seja, 126. (observação 1)

Neste excerto, apesar de Jaime ter tido algum cuidado em referir aspetos descritos no enunciado, o enfoque da sua explicação esteve no procedimento e no valor numérico obtido.

Numa outra tarefa, *Ração gasta num dia* (t6), inserida na *ficha desafio*, estavam incluídos 15 dados numéricos contínuos e requeria-se que estes dados fossem agrupados em determinadas classes fornecidas no enunciado, e que fosse calculada a média, entre outras questões. Quando, na apresentação do grupo, é mostrado o cálculo da média, Estela estava na expetativa que este valor fosse determinado tendo em conta a nova organização dos dados (tabela com os dados agrupados em classes, que incluía uma coluna com as frequências absolutas de cada classe) ou seja, que se usasse o valor da marca de cada classe para realizar o cálculo. A professora ficou muito surpreendida pelo facto de os alunos terem recorrido aos 15 dados originais para calcular a média e de não se terem apercebido da possibilidade de proceder ao cálculo da média usando os dados da tabela preenchida, nomeadamente, a marca da classe. Esta ocorrência pareceu dever-se ao facto de os alunos do grupo não se lembrarem ou não saberem como calcular a média para o caso dos dados agrupados em classes. Este assunto não tinha sido revisto na aula, nem mencionado pelos alunos ou pela professora antes desta situação. Mostra-se, de seguida, a interação da professora com os alunos do grupo:

> **Clara (grupo):** Por fim, a última alínea pede para calcular a média, isso não é novidade nenhuma!… É só somar os valores que temos aqui em cima (apontando para os dados originais) e dividi-lo pelo número de dias.
> **Professora:** (…) Outra coisa… calcularam a média à moda antiga. E à moda nova… partindo de dados agrupados?
> **David:** À moda nova?
> **Clara (grupo):** Não sei.
> **Professora:** Quem sabe?... O grupo que vem a seguir vai explicar como é que se calculada a média em dados agrupados?
> **Diana:** Na calculadora, Stora?

> **Professora:** Sim, e em dados agrupados?
> **Diana:** Eu não me responsabilizo… em dados agrupados?... Não sei se os nossos dados são agrupados. (observação 1)

Pelo facto de os alunos terem acesso a todos os dados numéricos (aos 15 dados), a opção de calcularem a média com base nessa informação parece ter sido considerada a mais natural para eles, levando-os a não sentirem necessidade de recorrer à tabela e, por conseguinte, a realizar uma estimativa desse valor. O resultado foi apresentado mas não interpretado à luz do contexto dos dados. A professora também deu a entender que estava à espera que os alunos do grupo se deparassem com a média para dados agrupados em classes enquanto exploraram este assunto, na resolução da tarefa e preparação da sua apresentação. O excerto evidencia que havia alunos que não tinham a perceção do que significava os dados estarem agrupados.

Quando se voltou a focar em dados agrupados em classes, Estela optou por usar o software *e-escola*[20] da Porto Editora. Neste contexto, no que concerne à média, referiu que esta poderia ser calculada através da fórmula usual, embora usando os valores da marca de cada classe e respetiva frequência, mencionando, também, que este cálculo poderia ser facilmente obtido na calculadora gráfica: "depois de se determinar a marca de cada classe, a calculadora gráfica usa-se tal e qual como se faz para a variável quantitativa discreta. Colocam na *lista 1* as marcas das classes e na *lista 2* as suas respetivas frequências e usam o *1-Vars-Stats*, como habitualmente" (observação 4).

Apresenta-se, de seguida, a interação que houve entre professora e alunos na única tarefa, desenvolvida na aula, que solicitava a realização de associações entre dois valores numéricos de médias (16 ou 16,3) e diagramas de barras (A, B, C e D, figura 22), designada por Tarefa *Distribuições das idades das equipas de trabalho*[21] *(t10)*:

---

[20] A Escola virtual foi adquirida pela escola e encontrava-se instalada em vários computadores da escola, em particular nos computadores das salas de aula. Com os dados agrupados, depois de ter projetado o *e-escola* sobre a tela branca, Estela apresentou a regra mais conhecida para definição de classes $2^k \geq n$ (em que o número de classes, *k*, a considerar numa amostra de dimensão *n*, é o menor inteiro que satisfaz a fórmula), apenas para os alunos tomarem conhecimento. De seguida, selecionou uma situação já com dados agrupados em classes e indicou a marca da classe de cada uma delas em interação com os alunos.

[21] De referir que na tarefa *Distribuições das idades das equipas de trabalho*, o gráfico A correspondia a uma distribuição em que as barras tinham todas a mesma altura e o gráfico D, por sua vez, era um gráfico de barras simétrico (com alturas de barras crescentes até à barra central, e com as seguintes decrescentes). Estas duas distribuições diziam respeito a situações paradigmáticas em que o valor da média equivaleria

Professora: Ora bem, vamos lá ver se entendemos o que eu pretendo que façam… Temos aqui quatro, quatro quê?
Vários alunos: Gráficos de barras…
Professora: Ah…, quatro gráficos de barras, o A, o B, o C e o D, que representam as distribuições das idades dos elementos que constituem cada um dos quatro grupos de trabalho de uma colónia de férias. E depois temos aqui para cada um dos grupos, não sabemos qual é qual, estes pares de valores. Isto aqui é? [apontando para a notação da média - $\bar{x}$]
Vários alunos: A média.
Professora: O outro? [apontando para a notação s]
Vários alunos: O desvio-padrão.
Professora: O que pretendo que vocês façam… têm aqui 4 hipóteses e quero que vocês façam corresponder cada uma delas a cada gráfico…Pensem, não desatem já a dizer… Comecemos pela média, está bem?
Professora: (passado uns minutos, pergunta) Já descobriram qual é a média de algum gráfico?
Alguns alunos: É 16 [referindo-se à média do gráfico A]
Professora: Para aqui a média é 16. Onde é que também é 16?
Alguns alunos: Na D. (observação 5)

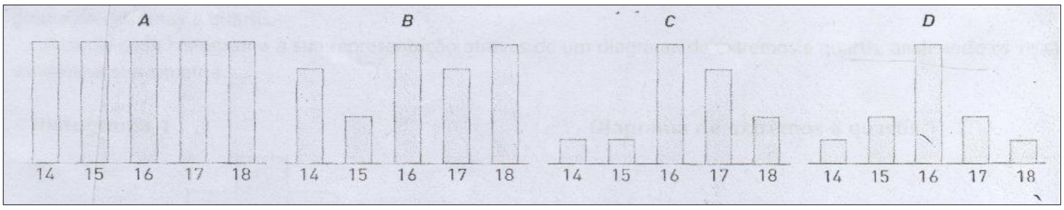


Figura 22: Diagramas de barras.

Na aula, assim que ficou assumido que os gráficos A e D corresponderiam à média 16, os outros dois foram automaticamente referidos como os que teriam de ter obrigatoriamente média 16,3. Não houve nenhuma intervenção, nem de alunos, nem da professora, justificando o facto de os gráficos B e C terem de ter uma média superior aos de A e D.

Observou-se, no excerto acima, que os alunos identificaram corretamente quais as distribuições que tinham média 16 (gráfico de barras A e D). Contudo, eles nunca foram questionados a justificar as suas respostas, não fizeram qualquer comentário às

ao valor da idade da barra central dos gráficos (idade 16). Já os gráficos B e C correspondiam a distribuições assimétricas e apontar a estimativa correta do valor da média para os gráficos B e C poderia não ser tão imediato como para as outras duas situações com os gráficos A e D. As distribuições apresentadas nos gráficos B e C tinham em comum a barra central com a mesma altura (que corresponde à idade 16) e ainda o facto de terem mais dados à direita desta barra central. Esta observação poderia ser particularmente útil para se demonstrar que a média destas distribuições teria de ser superior a 16 anos.

diferentes distribuições gráficas que tinham em mãos e também não levantaram qualquer dúvida ou questão relativamente às associações que assumiram.

Moda

Na aula, a moda era habitualmente identificada pelo dado ou dados com maior frequência. Como exemplo, mostra-se a resposta dada por um aluno, Jaime, na primeira tarefa que a envolvia, tarefa *Numero de hóspedes de uma estância balnear* (t7):

> (…) como todos sabem a moda é o acontecimento que ocorre mais vezes, podemos ver aqui que a moda é 80 [surge duas vezes na lista dos sete dados do enunciado], todos os outros [dados] só aparecem uma vez. (observação 1)

Através do *e-escola* (observação 4), Estela mostrou um exemplo de um conjunto de dados bimodal e fez referência à existência de distribuições com mais do que duas modas (plurimodal) ou eventualmente sem moda (amodal). Também apresentou uma tabela de frequências absolutas com dados agrupados em classes e um histograma (ver figura 23), nos quais destacou a classe modal. Contudo, a referência à moda quando foram analisadas representações tabulares e gráficas, no decorrer das aulas, não mais emergiu na interação e nas discussões entre professora e alunos.

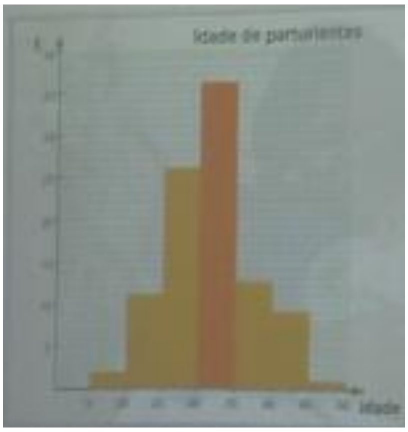


Figura 23: Histograma com a barra que corresponde à classe modal destacada.

Mediana e quartis

Este subtópico foi sobretudo desenvolvido e também apresentado na sala de aula com dados não agrupados em classes, nomeadamente, através das tarefas t7 e t4 e da situação *Mediana na tabela* (anexo 3). A professora também propôs uma outra tarefa (t15) que versava a mediana em representações gráficas. Embora não tivesse proposto nenhuma tarefa que englobasse a identificação/determinação da mediana e restantes quartis (primeiro e terceiro) com dados agrupados em classes, não deixou de referir a situação *Quartis em gráficos* (anexo 3) através do *e-escola*, em que mediana, classe mediana e quartis foram graficamente identificados através de um histograma de frequências relativas acumuladas e respetivo polígono de frequências.

Tarefa *Número de hóspedes de uma estância balnear* (t7) e Situação *Mediana na tabela* (anexo 3)

Nesta tarefa, o conceito de mediana foi inicialmente explicado através de um procedimento a seguir, que consistia em observar uma lista ordenada de dados e associar consecutivamente os seus extremos um ao outro, até se alcançar o valor central (caso a lista tivesse um número ímpar de dados) ou os dois valores centrais (caso a lista tivesse um número par de dados):

> **Jaime (grupo):** Para a mediana colocámos os números todos alinhados [ordenadamente]…e, se repararmos, há um ali (apontando para o dado na extremidade direita), outro aqui (apontando para o dado na extremidade esquerda)… depois um ali, outro aqui (apontando para as extremidades interiores seguintes e procedendo desta forma consecutivamente)… verificámos que no centro encontrámos o 100.
> **Lurdes:** David, e se for par? [o número de dados]
> **Jaime (grupo):** Os dois números que ficassem a meio… Calculavas a média deles. Como é impar… [a mediana] é o 100, [assim] consegues descobrir a mediana.
> **Lurdes:** Ah, já percebi. (observação 1)

Nesta interação destaca-se também o à vontade da aluna Lurdes em questionar o colega durante a sua apresentação.

A mediana foi localizada numa tabela de frequências absolutas acumuladas (ver Situação *Mediana na tabela* (anexo 3) e figura 24), que dizia respeito ao tamanho do calçado que 115 mulheres calçavam (a variável assumia valores de 35 a 39, inclusive, e ainda, o 41 e o 42). Mostra-se a interação que houve na sala de aula com vista à descoberta do valor e da localização da mediana na tabela:

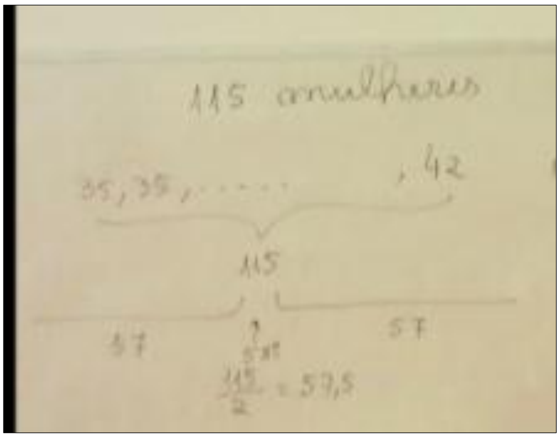


Figura 24: Registos de localização da mediana (no quadro) e na tabela de frequências.

> **Professora:** Nesta tabela os tamanhos dos sapatos já estão organizados por ordem crescente. Quantas mulheres eram no total?
> **Alguns alunos:** 115.
> **Professora:** Para calcular a mediana tenho de ter estes valores listados… a começar no 35 que tem de estar duas vezes (lendo esta frequência na tabela) até ao 42 que está uma vez… ao todo são 115 mulheres (regista no quadro valores – ver figura acima). Então qual a posição central? Como faço para descobrir quem fica no meio?... Faço 115/2… Que dá?
> **Mário:** Mais ou menos 58.
> **Ivo:** Dá 57,5.
> **Professora:** Isto quer dizer que há uma pessoa no meio ou não há?
> **Ivo:** Há a 58!
> **Professora:** Exato! Vai haver vai!… Vai haver 57 para um lado, 57 para o outro e no meio está quem ocupa o quinquagésimo oitavo lugar. Agora na tabela… onde é que é mais fácil ver?… Na coluna das frequências acumuladas… procuramos a 58.º lugar. Onde é que ele está? No 38?
> **Alguns alunos:** Não.
> **Professora:** É no 68? (ver figura 24)
> **Alguns alunos:** Sim.
> **Professora:** É aqui. É alguém que tem tamanho 38 [o valor mediano].
> **Mariana:** Quer dizer que a mediana é 38, não é?
> **Professora:** Exato. Está?! Perceberam? Pronto. (observação 4)

Esta situação foi apenas dedicada à mediana. A, professora ao interagir com os alunos, determina a posição da mediana após o cálculo de metade do tamanho da amostra e identifica-a na tabela através das frequências absolutas acumuladas.

Além de ter sido explicitamente identificada como o valor que divide a distribuição de dados em duas partes iguais, a mediana foi também reconhecida como o segundo quartil. Na tarefa *Número de hóspedes de uma estância balnear* (t7), o procedimento usado para se obter a mediana foi o referido para se alcançar os valores dos 1.º e 3.º quartis:

> **Diana (grupo):** (Para introduzir os quartis disse) Também podemos considerar uma distribuição dividida em quatro partes. Temos aqui os nossos números [referindo-se aos dados] vamos dividi-los em quatro quartos. O nosso mínimo é 80 e o nosso máximo é 240. Eles já estão por ordem crescente [mostra a lista de dos 7 dados com os quartis assinalados]. Ao dividirmos temos o primeiro quartil [apontando para o segundo dado de valor 80 na lista] (…). O primeiro quartil [80] separa os primeiros 25% dos dados dos restantes 75%, já o terceiro quartil [180] separa os primeiros 75% … que são maiores ou iguais ao quartil… dos outros 25% que são maiores ou iguais ao quartil. O segundo quartil [100] é a mediana. Perceberam? (Ouve-se vários risos)
> Vou voltar a repetir… Então nós temos os nossos números [referindo-se aos dados] ordenados por ordem crescente, vamos dividi-los em duas partes, como foi com média [há alguns alunos que dizem "mediana" corrigindo Diana] como na mediana como o David mostrou [de valor 100] e [agora] cada uma delas vai ser dividida em duas partes, temos 4 partes, certo? (…) … 100% a dividir por 4 dá 25%, certo? Cada uma das partes corresponde a 25%... Ao dividir… o primeiro quartil separa os primeiros 25% dos restantes 75%. Está certo?
> O terceiro quartil separa os primeiros 75% menores ao igual ao quartil dos restantes 25% que são maiores ou iguais ao quartil. O segundo quartil como já vimos é a mediana, certo?
> Toda a gente percebeu? (ninguém responde) Percebeste Cátia? (ouvem-se risos)
> **Professora:** Ela [Diana] tem razão, a Cátia estava distraída… Diana, posso só dar um exemplo? Tenho um bolo… quero a quarta parte do bolo. O que tenho de fazer ao bolo?
> **Vários alunos:** Dividir em quatro partes.
> **Professora:** Exatamente, dividir em quatro… e cada uma delas é um quarto. Era isso que a Diana estava a dizer. (…) No fundo, quartil vem de dividir em quarto… (a professora identifica no quadro cada um dos quartos, cada conjunto que corresponde a 50% dos dados e cada conjunto que corresponde a 75% dos dados através da lista dos dados ordenados com os valores dos quartis assinalados)… Na totalidade são quantos por cento?
> **Vários alunos:** 100%
> **Diana (grupo):** Isto dos quartis é principalmente utilizado com grandes quantidades de dados. Aqui como estamos a aprender é mais fácil assim... (…) e [com os quartis] dá para ter uma ideia geral da localização dos resultados (referindo-se aos dados). (observação 1)

A aluna Diana, na primeira vez que explica por palavras suas o que são os quartis, engana-se na descrição do terceiro quartil, mas, na segunda vez em que repete essa explicação, faz um comentário descritivo correto referente ao terceiro quartil. Esta aluna teve também necessidade de voltar a explicar a correspondência entre quartil e respetiva percentagem de dados, por não estar a receber *feedback* dos colegas. Já por causa disso, a professora interveio para lhe dar uma ajuda. Diana considerou os quartis como medidas de localização que dão uma perspetiva geral de como os dados estão distribuídos, transmitindo, deste modo, uma ideia informal de variabilidade.

Na resolução da tarefa t7, apesar de se ter feito referência ao facto de o método para se identificar o primeiro e terceiro quartis ser equivalente ao que se fazia para a mediana, não ficou claro se os subconjuntos de dados que serviam de base para a identificação do primeiro ou terceiro quartis incluíam ou não o valor da mediana. Efetivamente, na situação concreta, com os sete dados desta tarefa, tendo em conta a resolução apresentada pelo grupo, a mediana não era incluída nos ditos subconjuntos. Contudo, este pormenor importante não foi evidenciado na interação que houve entre alunos e professora na correção da tarefa.

Ainda no desenvolvimento da tarefa t7, foi construído o diagrama de extremos e quartis e indicou-se como poderia ser visualizado na calculadora gráfica. A construção do diagrama de extremos e quartis (o que inclui os dois "bigodes" e a caixa) foi apresentada por outros dois elementos do grupo (Pedro e Sara). O Pedro começou por alertar que a "escala deve abranger todos os nossos números (referindo-se aos dados)". Quando mostrou o diagrama de extremos e quartis já feito no acetato projetado, houve um aumento do nível de som na sala, provavelmente fruto dos alunos não estarem a entender o gráfico (ver figura 25):

> **Pedro (grupo):** Isto aqui é o mínimo (apontando para o valor 80) e aqui está o máximo que é o 240.
> **Mário:** Aquele traço na vertical é um quartil? (referindo-se ao traço que se encontra na extremidade esquerda do gráfico)
> **Sara (grupo):** Sim. É assim… Aponta para a localização dos traços [verticais pequenos] do 1, 2 e 3 quartis. Percebeste? … O traço [vertical] maior corresponde ao mínimo e ao máximo. Nós vimos que o mínimo é 80 mas também dá para corresponder ao [primeiro] quartil…. É assim… o [primeiro] quartil… é como… se… houvesse um tamanho errado, o primeiro quartil é igual a 80. Por isso é que não está o espaço tão distanciado… como entre o 3.º quartil e o máximo… (observação 2)

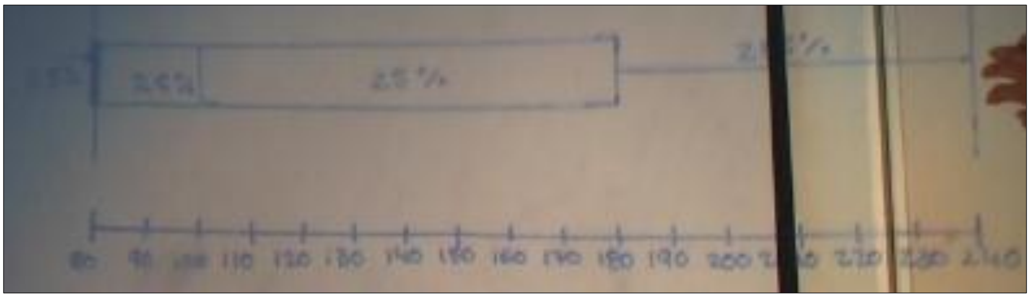


Figura 25: Diagrama de extremos e quartis (projetado na aula).

Pedro identificou o mínimo e o máximo da distribuição no diagrama de extremos e quartis que apresentou. Mário acabou por fazer uma pergunta pertinente ao reparar que o mínimo e o primeiro quartil tinham o mesmo valor. Sara explicou que este gráfico não tinha o bigode do lado esquerdo precisamente por causa dos valores do mínimo e primeiro quartil coincidirem, acabando por referir-se à ausência do bigode esquerdo até de uma forma curiosa, como se existisse um "tamanho errado" entre mínimo e o primeiro quartil. Graficamente, tiveram o cuidado de representar o mínimo com um segmento vertical mais comprido e o primeiro quartil com um segmento vertical mais pequeno e mais grosso do que o anterior.

A discussão continuou com o apoio da calculadora gráfica, tal como a última alínea desta tarefa solicitava:

> **Professora:** Vamos ver como fizeram na máquina (gráfica)?
> [alunos do grupo dirigem-se para o computador que está na secretária da professora]
> **Sara (grupo):** (começa por inserir os 7 dados na lista 1 *L1*) Posso? Já sabem como é a lista… Há uma maneira de obter a média, os quartis de uma vez. Vão a *Calc* e selecionam *1-VarStats* e colocam a lista que querem… O *x barra* corresponde à média, que é… aproximadamente 126.
> **Professora:** Foi o que vos deu à mão, não foi?
> **Sara (grupo):** Sim. O *n*=7 dá o número de variáveis que temos na lista.
> **Professora:** (corrigindo a Sara) Dá o número de dados!
> **Sara (grupo):** Até temos o mínimo que é 80… o primeiro quartil que também é 80 e a mediana dá 100 que também é o segundo quartil e o terceiro quartil que é 180.
> **Professora:** Anda mais para baixo no ecrã [para se poder ver o valor do máximo].
> **Sara (grupo):** E o máximo é 240. Assim retiram toda a informação sem fazer nenhum cálculo à mão! Para obter o diagrama de extremos e quartis, basta ir a *statplot 1*, selecionar o gráfico que se pretende… na *Xlist* colocar *L1* e na frequência (*Freq*) colocar 1… a frequência de cada dado é 1 (à frente *de Freq* a calculadora tinha registado *L2*; embora Sara soubesse que em vez de *L2* deveria

estar o valor 1, já não se lembrava como fazer essa alteração, nem os seus colegas do grupo, a aluna acabou por ser ajudada pela professora).
Professora: Clica na tecla alfa e escreve 1. Arranja a janela (menu *window*). Isso! [ao ver a aluna a ajustar os valores *Xmínimo* e *Xmáximo*]... Já está a aparecer! [o gráfico]. Coloca o ecrã grande (ver figura 26)… Fantástico! (observação 2)

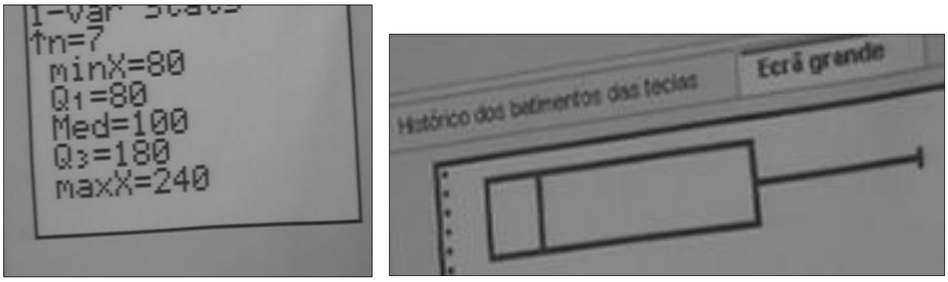


Figura 26: Ecrãs da calculadora gráfica com os valores dos quartis e respetivo diagrama de extremos e quartis.

Sara mostrou aos alunos quais as instruções a seguir na calculadora gráfica para chegar-se às medidas estatísticas tendo em conta que, neste recurso, há várias determinadas em simultâneo: média, mínimo, máximo e quartis da distribuição. As explicações fornecidas por Sara tinham também em vista mostrar aos colegas como representar o diagrama de extremos e quartis na calculadora para os dados da tarefa. Ajudada pela professora, esta aluna acabou por ter que ajustar alguns parâmetros para obter o gráfico, nomeadamente, na frequência escrever 1 (porque não havia uma lista adicional com a frequência dos dados) e ainda os parâmetros do *Zoom* para ser possível observar o gráfico no ecrã da calculadora projetada. Verifica-se que não foi realizada uma leitura deste gráfico tendo em conta os dados e o seu contexto.

Tarefa *Classificações a Matemática das turmas A e B* (t4)

Esta tarefa, que tinha sido usada no tópico anterior (Organização de dados), foi também explorada neste tópico quando Estela pediu ao grupo de alunos que a tinha apresentado na aula, para irem novamente ao computador da sala de aula mostrar à turma como é que se fazia o diagrama de extremos e quartis na *TI-Nspire.* Assim que o diagrama de extremos e quartis apareceu exibido na tela (ver figura 27), a professora dirigiu-se à turma e pediu que se fizesse uma análise aos resultados da turma A desta tarefa:

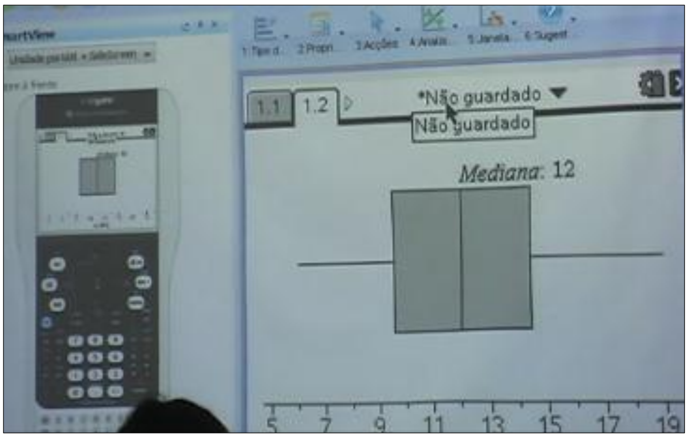


Figura 27: Diagrama de extremos e quartis das notas da turma A.

Professora: (…) Que análise é que podem fazer desta turma, que era a turma…
Alunos: A
Professora: (…) O que é que podem dizer? Podem pedir que o Leonardo ande com o rato para onde quiserem [de modo a verem os valores dos quartis]. Qual é o valor do primeiro quartil?
Ivo: 9,5
Professora: O que acham desta turma?
Ivo: É jeitosa.
Professora: Porque? É jeitosa porque…
Dalila: Os primeiros 25%...
Professora: …Das notas são? Quem ajuda? O que podem dizer? Diz?
Ivo: os 25% dos piores alunos, a média das notas deles é 9,5.
Professora: não sabes se é a média! Mas sabes quantos por centos existem de negativas.
Ivo: 25%
Professora: Porque?
Dalila: Os outros 75% é de 9,5 para cima, que já é positivo.
Professora: Exatamente. 9,5 já é 10 valores. Então 75% nesta turma são positivas. Entre que valores estão 25% das notas maiores desta turma [entre Q3 e o máximo]?
Jaime: Entre 14,5 e 19.
Professora: E 50% das notas positivas estão entre que valores (referindo-se à caixa)?
Jaime: 12 e 19.
Professora: Podia ser… E entre 12 e 14, 5 [entre Q2 e Q3]…
Jaime: Aí é 25%.
Professora: Pronto, perceberam como é que faziam? Isto só se adquire com prática.
Jaime: E com uma máquina dessas!!
Professora: Mas agora também pegamos na outra [*TI 83* ou *84 Plus*]. Com a outra também já fizeram um diagrama de extremos e quartis, eu só não sei fazer o diagrama circular nela [*TI 83* ou *84 Plus*] mas também é possível. (observação 2)

Esta foi a primeira vez foi feita uma leitura a um diagrama de extremos que tivesse em conta o contexto dos dados. Inicialmente, o aluno Ivo interpretou erradamente o primeiro quartil como uma média dos primeiros 25% dos dados mas Estela corrigiu-o imediatamente. A resposta pretendida pela professora era que os alunos chegassem ao entendimento de que cerca de 75% das notas da turma eram positivas. Depois de se ter chegado a essa resposta, houve uma interação em torno da identificação de outros intervalos no diagrama de extremos e quartis, nomeadamente, os de 25% e 50% de dados. Nas suas respostas, os alunos usaram os valores dos quartis como referência. Estela também disse aos alunos que as leituras que tinham acabado de fazer exigiam algum treino.

Tarefa *Associa entre histogramas e diagramas de extremos e quartis* (t15) *e* Situação *Quartis em gráficos* (anexo 3)

O enunciado desta tarefa incluía oito gráficos (uma coluna de quatro histogramas e outra de quatro diagramas de extremos e quartis – ver figura 28). A professora solicitou que os alunos analisassem essas representações e identificassem as que correspondiam à mesma distribuição. Quando os alunos estavam a explorar esta tarefa, Estela sugeriu também que usassem a "exclusão por partes" (observação 5) como estratégia. Depois de dar algum tempo aos alunos para pensar na tarefa, a professora pediu que partilhassem as suas respostas. Eis aqui a interação que houve entre ela e os alunos respondentes:

> **Ivo:** Acho que o Histograma 1 é o Diagrama [de extremos e quartis] 4.
> (Vários alunos expressaram que concordavam o que o Ivo disse)
> **Professora:** Porquê, explica? (Enquanto assenta no quadro a correspondência entre histograma 1 e diagrama 4).
> **Ivo:** O histograma 1 é o único simétrico e o diagrama 4 também.
> **Professora:** Não é parvo o que estás a dizer!... Como é simétrico, achas que a mediana está onde (apontando para o histograma 1 que está projetado sobre o quadro branco)?
> **Ivo:** No…na barra do centro.
> **Professora:** Exatamente. Mas eu também podia pensar que era o diagrama de extremos e quartis 1 [em vez do diagrama 4], não?
> **Ivo:** Não.
> **Professora:** Explica então…
> **Ivo:** Então se fosse o de cima (diagrama de extremos e quartis 1) para já não é simétrico, é mais virado para a esquerda… e as barras nos extremos (referindo-se aos "bigodes" do diagrama) teriam de ser maiores para o intervalo dos 25%.
> **Professora:** Tem de se ver simetria nos dois gráficos, certo?... Quem fala de outros?

David: Acho que o último histograma (histograma 4) corresponde ao primeiro (diagrama 1).
Professora: Porquê? (enquanto assenta no quadro a correspondência entre histograma 4 e diagrama 1)
David: (Referindo-se ao histograma 4) Temos uma barra a meio, não é?… de um lado e do outro são praticamente simétricos… só que as barras de um dos lados [direito] são mais altas.
Professora: Exatamente. Aquela mínima diferença…
David: Faz toda a diferença!
Professora: Exato. Então achas que a mediana está mais ou menos aonde?
David: A meio…
Professora: Portanto, dos diagramas de dispersão que sobram só temos aqui um, que é o primeiro (apontando para a largura da caixa entre o 2.º e 3.º quartis do diagrama de extremos e quartis 1). Isso foi provocado por quem?
Dalila: Aquela barrinha do meio…
Professora: Pela tal… aquela barrinha do meio mais alta (apontando para a quinta barra no histograma 4) bastou para provocar uma dispersão maior (no diagrama). (observação 5)

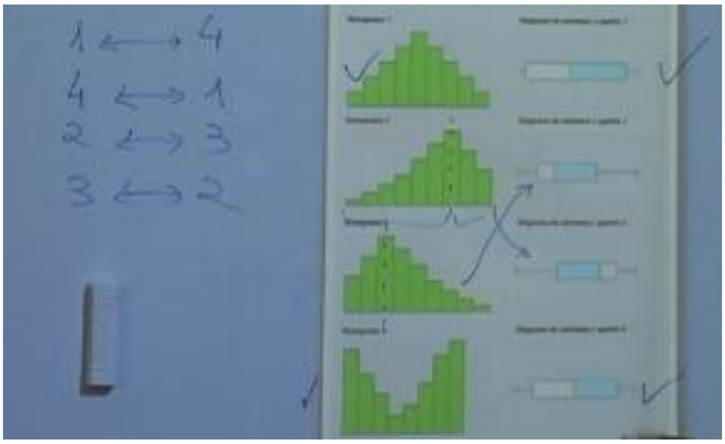


Figura 28: Enunciado da tarefa e as associações entre histogramas e diagramas (respetivamente) assinaladas no quadro.

Na continuação desta discussão, Estela foi pedindo aos alunos que justificassem as suas opções, levando-os a fazerem uma análise das representações que não tivesse apenas em conta a questão da simetria, questionando-os sobre a localização da mediana nas duas representações gráficas e a sua consistência. Sobre este assunto, há, inclusivamente, um aluno (Leonardo) que vai ao quadro marcar a tracejado uma estimativa da localização da mediana nos histogramas 2 e 3 (ver figura 28). Estela afirmou concordar com os tracejados marcados pelo aluno. É de referir, também, que os restantes histogramas (2 e 3) foram corretamente e facilmente identificados pelos alunos que chegaram à conclusão que correspondiam a distribuições enviesadas, uma à direita e a outra à esquerda, e que os associaram ao diagrama 3 e 2, respetivamente.

No excerto acima, o aluno Ivo indicou oralmente a localização da mediana no histograma 1, de modo acertado, e verificou que esta não correspondia à mediana assinalada no diagrama 1. Além disso, no seu raciocínio, teve também em consideração a variação dos dados representados no diagrama 1, especificamente nos seus primeiros e últimos 25%. Ao notar a não simetria deste diagrama 1, eliminou, assim, a hipótese do histograma 1 lhe corresponder. Contudo, este aluno justificou a não simetria que observou no diagrama 1 de uma forma não totalmente clara, dizendo: "é mais virado para a esquerda" (observação 5). Embora não tendo questionado este aluno sobre o que ele queria dizer, pela sua intervenção, a professora parece concordar que o diagrama 1 não é efetivamente simétrico. Um outro aluno, David, assumiu que o histograma 4 só poderia corresponder ao diagrama 1 por se relacionarem ambos com distribuições quase simétricas. No entanto, a discussão entre os alunos e a professora não possibilitou que fosse detetado o erro que havia no enunciado tarefa, e que era que o diagrama 1 não poderia corresponder ao histograma 4, como era suposto acontecer.

Estela não se apercebeu que a localização da mediana no histograma 4 no centro da distribuição, tal como o David apontou, não era a mais adequada (deveria ser antes ligeiramente à direita desta indicação), o que poderia ter ajudado a revelar que as distribuições apresentadas no histograma 4 e no diagrama 1 não poderiam corresponder à mesma. Além disso, à semelhança da aluna Dalila, Estela, ao considerar que havia uma "dispersão maior" (observação 5) no histograma 4 pelo motivo da quinta barra ser mais alta do que a quarta, acabou por transmitir a conceção errónea de associação entre maior densidade nas barras (nomeadamente, na metade direita do histograma) com a maior amplitude observada entre o 2.º e 3.º quartis do diagrama de extremos e quartis 1. Esta leitura no histograma 4, na perspetiva da professora, estaria em conformidade com o facto do diagrama de extremos e quartis 1 apresentar um ligeiro enviesamento para a direita, ditando a correspondência entre estas duas representações. De uma maneira geral, esta tarefa permitiu que fossem evidenciadas algumas características comuns dos dados e sua variabilidade entre as representações gráficas associadas.

Estela selecionou do *e-escola* um conjunto de dados agrupados em classes, através da qual apresentou a Situação *Quartis em gráficos* (anexo 3). Esta situação expunha algumas representações associadas a estes dados, nomeadamente, o histograma das frequências absolutas, o histograma das frequências absolutas acumuladas, e o histograma das frequências relativas acumuladas conjuntamente com o respetivo polígono (de frequências relativas acumuladas). Com este último gráfico, que combinava duas representações, Estela explicou uma forma possível de marcar no histograma os quartis (Q1, Q2 e Q3) com o apoio do polígono: "ao encontrar a abcissa do ponto que correspondia, respetivamente, a uma ordenada de 25%, 50% e 75%" (observação 4). Identificou também a classe mediana e associou o diagrama de extremos e quartis ao polígono/histograma das frequências relativas dos mesmos dados agrupados (ver figura 29):

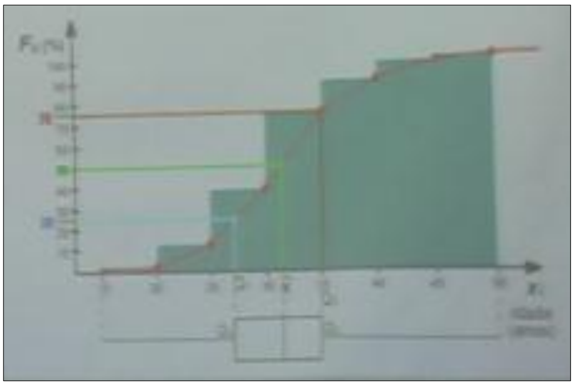


Figura 29: Relação entre o polígono de frequências e o diagrama de extremos e quartis.

Depois desta apresentação da professora, o aluno Mário interveio, parecendo um pouco confuso, questionando a professora sobre diferentes formas de determinar a mediana em dados agrupados em classes:

> **Mário:** Oh Stora… Só se pode saber a mediana… no gráfico das frequências acumuladas [referindo-se ao histograma]?!
> **Professora:** Se quiseres localizá-la graficamente, sim. Mas também a calculas na máquina… Como é que a fazes na máquina, depois de veres a marca de cada classe?
> **Mário:** Mas manualmente, só dá assim?...
> **Professora:** Sim, só assim, no gráfico. Ou então fazes na calculadora como fazias para os dados discretos para também a obteres manualmente. (observação 4)

A professora esclareceu o aluno Mário que era possível localizar graficamente a mediana, tal e qual como tinha acabado de mostrar com o histograma e polígono de frequências absolutas acumuladas. Também referiu que era viável calculá-la, sugerindo uma transformação nos dados agrupados, ou seja, que cada classe fosse substituída pela sua marca e que se processasse o cálculo da mediana da mesma forma como se fazia com os dados discretos. Porém, não foi dito que com qualquer um dos métodos referidos se obtém uma estimativa do valor da mediana. A reação do aluno não pareceu de todo surpreendente, visto esta ter sido a única experiência na aula com mediana em dados agrupados em classes.

5.3.3.2 Reflexões sobre o tópico na aula

Na entrevista dois, Estela é perentória quando refere em que consiste essencialmente a aprendizagem das medidas de localização (média, moda e mediana): "saber determiná-las através de cálculos", acrescentando que os alunos devem também tentar "visualizar ou estimar essas medidas em gráficos de barras ou histogramas" sempre que lhes for solicitado. Na sua perspetiva, o cálculo é acessível, e o trabalhar as medidas nas representações gráficas constitui um desafio provavelmente menos alcançável para os alunos.

Para Estela, a noção de mediana que os alunos reconhecem é de que "50% [dos dados] é para um lado e 50% [dos restantes dados] para o outro" (entrevista 3). Com a tarefa *Associa entre histogramas e diagramas de extremos e quartis* (t15) os alunos, a pedido da professora, tentaram, em particular, localizar a mediana nos quatro histogramas e averiguar a sua consistência nos respetivos diagramas de extremos e quartis. Estela indicou a vantagem desta estratégia usada: "Onde é que está a mediana em cada histograma? É uma ajuda [para se resolver a questão] … imaginar visualmente onde estará a mediana nos gráficos." (observação 5). Estela considera esta situação que lhes colocou inovadora, pois: "[analisar a mediana] visualmente era uma coisa que não se fazia, era tudo feito teoricamente. Nada de representações… nunca se usava isso" (entrevista 3). Mencionou ainda que encontrou esta tarefa num manual recente (não no adotado pela escola): "pela primeira vez me dei a pensar como abordar estes gráficos… e acho que os alunos viram muito bem as ligações quando os discutimos" (entrevista 3), acrescentando que pesou no facto de a ter trazido à aula ter participado num trabalho semelhante realizado na formação de Estatística:

> … e claro que fiquei mais sensibilizada com os teus diagramas da formação. Lembras-te que puseste aqueles histogramas? E os colegas [da formação] começavam a comentar [sobre a localização da média e da mediana] e eu ficava ali a olhar… 'como é que eles vêm isso?!' (…) nunca me tinha desprendido (…) (entrevista 3).

Estela reconhece que, através da ação de formação, viu, pela primeira vez, ser discutido como procurar localizar a média e a mediana em histogramas, considerando esta atividade de interesse para o ensino da Estatística.

Na perspetiva da professora, os alunos parecem ter facilidade em identificar distribuições simétricas, distribuições aproximadamente simétricas e distribuições não simétricas ou enviesadas. Mas defende que, para os alunos, a média pode ser mais difícil de se aceder visualmente nalgumas representações (histogramas, gráfico de barras) do que, por exemplo, a mediana, explicando o motivo da seguinte forma:

> (…) porque a média é muito influenciada por valores extremos mas acho que eles [alunos] conseguiram ver mais ou menos se a média ficava ou à esquerda ou à direita da mediana (…) A ideia intuitiva é tentar aproximar a média da mediana, mas às vezes não é [possível], só quando a distribuição [de dados] é simétrica. (entrevista 3)

Neste excerto, a professora deixa transparecer o conhecimento que tem sobre a resistência da média face a valores extremos ou *outliers*, tratando-se de uma medida bastante sensível, mas que, curiosamente, foi um assunto que não emergiu de modo explícito na sua prática. Isto pode estar relacionado com o facto de, nas tarefas propostas ou mesmo na discussão de alguma delas, nunca ter sido ponderada a existência de dados discrepantes. Ainda recorrendo a este extrato, ao referir-se à visualização da média graficamente, Estela parece transmitir a ideia que, ao existirem valores extremos na distribuição de dados, estes afetarão a sua simetria. A visualização da média, por parte dos alunos, nesse caso poderá não ser tão fácil quanto no caso da mediana. Contudo, não tem dúvidas que os alunos não têm dificuldades em localizar a média e a mediana no caso da distribuição em causa ser simétrica.

O começar a trabalhar as medidas de localização graficamente permitiu que se refletisse sobre a posição de cada uma delas em distribuições enviesadas. Para uma distribuição enviesada para a direita, Estela referiu que a ordem das medidas de localização deveria ser primeiro a moda, depois a mediana, e depois a média tendo por base um esboço de diagrama de barras feito pela professora durante a entrevista. A professora situou a moda no pico e a mediana à sua direita; quanto à média disse:

> A média está depois da mediana por causa destes pequeninos (referindo-se às barras baixinhas de uma distribuição com cauda direita mais longa) mas isto sou eu agora a pensar depois de dar assim… de explorá-las em gráficos… eu [antes] nunca consegui ensinar isso… Mas [Sandra]… num 10.º ano de Matemática A… eu desafio-te a questionares algum professor ou colega que tenha só dado ensino regular… para ver se ele tem esta ideia ou se a decorou… coloca-lhe um exercício destes com gráficos de barras e pergunta onde estaria a mediana ou até a média… (entrevista 3)

Estela também crê que a maioria dos professores do ensino regular, que não tenha experiência na lecionação de MACS, não sabe localizar ordenadamente as medidas estatísticas em gráficos de barras ou histogramas.

A tarefa *Relatório sobre os resultados do teste intermédio*[22] (t12), que a professora teve o cuidado de apresentar na aula, acabou por ficar para trabalho de casa. Apenas 11 dos 25 alunos da turma entregaram este trabalho de casa. Oito relatórios mencionavam o diagrama de extremos e quartis; em três deles, os diagramas de extremos e quartis incluíam erros[23]. Dos oito alunos que apresentaram o diagrama de extremos e quartis corretamente, só um deles fez comentários ao diagrama: "A partir do gráfico acima, a nota mais baixa foi 5, a mais alta foi 18. O primeiro quartil é 7, o segundo quartil (mediana) é 10 e o terceiro quartil é 12. Como é possível observar a maior parte das notas situa-se entre 7 e 14, sendo entre 7 e 10 os resultados mais dispersos" (Comentário do aluno Leonardo ao diagrama de extremos e quartis que produziu para esta tarefa).

---

[22] Os relatórios produzidos pelos alunos foram recolhidos pela professora numa aula posterior à sua proposta.

[23] Cada um destes três relatórios vinha com pelo menos um dos quartis determinado erradamente e, assim sendo, estes diagramas não correspondiam à distribuição original de dados; em dois deles os quartis apareciam indevidamente assinalados nos seus diagramas que, em vez de corresponderem a um valor numérico, surgiam assinalados como distâncias.

Ao analisar esses relatórios dos alunos e, em particular, os diagramas de extremos e quartis produzidos a pedido da investigadora, a professora detetou os três diagramas de extremos e quartis errados. Ficou surpreendida com esses erros por achar que, como os alunos tinham à disposição a calculadora, não deveriam ter dificuldades em obter os valores dos quartis corretos de modo a elaborarem o gráfico certo. Ao examinar o comentário do aluno Leonardo ao seu diagrama de extremos e quartis teve a seguinte interação com a investigadora:

> **Professora:** Ele diz que "a nota mais baixa foi 5, a mais alta foi 18" – isto está certo?!... [continua a ler em voz alta] "O primeiro quartil é 7, o segundo é a mediana… é 10 e o terceiro quartil é 12. Como é possível observar a maior parte das notas situa-se entre 7 e 19" (Não concordando com o 19 assinalado diz) Não! … será entre 7 e 14?
> **Investigadora:** Sim, parece que está escrito 14.
> **Professora:** Então como é que ele sabe que é entre 7 e 14? Não é pelo gráfico!...
> **Investigadora:** Às tantas, ele deveria querer dizer entre 7 e 12, quererá referir-se à caixa?…
> **Professora:** Se fosse entre 7 e 12, não é a maior parte é 50%! … E agora… (continua a ler em voz alta) "(…) sendo entre 7 e 10 os resultados mais dispersos" … não acho!... Aqui ainda há maior dispersão (apontando para o bigode direito do diagrama de extremos e quartis) que vai entre 12 e 18, não é? (entrevista 2)

Nesta interação, a professora foi detetando incongruências entre a representação gráfica que o aluno fez e os seus comentários sobre ela. Tal como a professora disse, não lhe fazia sentido a referência do aluno ao intervalo [7, 14], dado que o 14 não era um quartil do gráfico. Por isso suspeitou que o aluno tivesse chegado ao 14 de outra forma, mas não pela observação do diagrama de extremos e quartis. Estela considerou que a variabilidade dos dados no diagrama de extremos e quartis foi erradamente descrita pelo aluno, mesmo que ele se tivesse enganado e trocado o 12 por 14. A professora é da opinião que, nessas circunstâncias, o intervalo entre o 1.º e 3.ºquartis não iria incluir a "maior parte" dos dados nem a maior dispersão dos dados indicada, pelo aluno, entre os 1.º e 2.º quartis.

Quando foram apresentadas à professora as respostas hipotéticas de alunos à análise de um diagrama de extremos e quartis *Respostas hipotéticas sobre um diagrama de extremos e quartis* (Entrevista 3, anexo 2), que correspondiam à distribuição das notas de uma turma de 24 alunos no último teste realizado de uma disciplina, a professora leu as afirmações calmamente, acentuando oralmente, nalgumas delas, certas partes (por exemplo, "Há exatamente", "Há pelo menos", "Existe muita dispersão", "Há uma grande concentração", "Há 25% de notas") dando a entender que careceriam de uma avaliação cuidada. Estela concordou com todas as respostas exceto com "Existe muita dispersão de dados". Na verdade, ao corrigir essa afirmação, considerou que há apenas "alguma dispersão entre 12 e 17 valores". Teceu também alguns comentários a algumas afirmações que considerou corretas, nomeadamente à afirmação: "a amplitude é de 13 valores", dizendo: "Exato… 17 menos 4 dá 13 de amplitude" e relativamente a "Há uma grande concentração de resultados entre 10 e 12 valores", dizendo: "Aliás [este intervalo] é o sítio onde há mais concentração nestas 4 partes ou… divisões… por quartis" (entrevista 3).

Os comentários feitos pela professora às afirmações foram, de uma maneira geral, corretos. Estela acentuou oralmente as expressões "há exatamente" e "há 25% de notas" mas não lhes fez nenhuma referência adicional, e avaliou os comentários que as envolviam como corretos. Contudo, a expressão "Há pelo menos…" seria a mais adequada nessas duas situações, visto que não se tinha acesso aos dados originais para se poder verificar a percentagem exata, apenas se tinha acesso às densidades nas diferentes partes do diagrama de extremos e quartis.

5.3.3.3 Discussão do tópico

CETE. Estela revela ter uma certa preocupação em promover atividades diversificadas na sala de aula, recorrendo a trabalhos de grupo e ao uso da tecnologia. Estela colocou diferentes grupos de alunos a trabalhar nas respetivas tarefas e a fazer apresentações à turma, que incluíam a explicação dos conteúdos envolvidos e a resolução da tarefa, interagindo com os restantes colegas. Adicionalmente, valorizou, na sua prática, o uso da *escola virtual* para referir mais representações associadas ao tópico (em particular, as que envolviam dados agregados em classes) e também para mostrar instruções sobre como usar a calculadora gráfica no cálculo das medidas estatísticas e na representação gráfica dos dados. Não foram propostas muitas tarefas nas aulas; contudo, elas eram diferentes umas das outras. Estela confessou que duas delas foram selecionadas e desenvolvidas pela primeira vez na sala de aula. Pensa-se que o facto de não ter detetado um erro no enunciado numa dessas tarefas, mesmo durante a sua correção em grande grupo, reflete a sua pouca experiência em lidar com o assunto envolvido. Na aula, surgiram alguns comentários e justificações (associadas à simetria e variabilidade dos dados nos gráficos) quanto à posição da mediana em representações gráficas, mas o mesmo não foi observado relativamente à moda e à média. Contudo, esses argumentos permitiram que algumas caraterísticas comuns e distintas entre essas representações fossem evidenciadas. Na tarefa *distribuições das idades das equipas de trabalho* (t10), Estela obteve as respostas corretas acerca da média, mas não questionou os alunos sobre os seus raciocínios. Na tarefa 7, obteve-se um diagrama de extremos e quartis peculiar, devido ao mínimo coincidir com o primeiro quartil. Na aula, não se fez uma leitura desse gráfico tendo em conta os dados e para se avaliar a sua utilidade na representação dos dados. As medidas de localização em dados agrupados em classes foram mencionadas de passagem, com uma situação específica.

CA. Durante a apresentação dos trabalhos de grupo relacionados com este tópico, cada aluno de cada grupo apresentou no mínimo uma resposta a uma alínea da tarefa que lhes foi atribuída. Estela revelou estar bastante satisfeita, de uma maneira geral, com a prestação de cada grupo, mesmo quando teve necessidade de intervir para apoiar uma aluna enquanto esta explicava o processo de identificação dos quartis (de um conjunto de dados quantitativos discretos) ou quando ajudou uma outra aluna a ajustar alguns parâmetros na calculadora gráfica para que chegasse à representação pretendida (um diagrama de extremos e quartis). A atuação de Estela na sala de aula tinha em vista promover a aprendizagem do aluno focada na exploração das situações através da calculadora gráfica (mesmo na mais avançada) quando poucos alunos a possuíam. Mesmo assim, o acesso às ideias e explicações dos alunos nas aulas, nas interações com vista à correção das tarefas neste tópico, foi geralmente limitado. Por exemplo, durante as apresentações dos dois grupos houve pouca participação dos restantes alunos da turma. Também na tarefa que ficou para trabalho de casa, *Relatório sobre os resultados do teste intermédio* (t12), Estela ficou um pouco dececionada com o facto de haver alunos a cometer erros no cálculo dos quartis e, consequentemente, na representação do diagrama de extremos e quartis.

CE. A professora revelou que o facto de ter trabalhado na ação de formação de Estatística a visualização e a estimativa do valor de medidas de localização em representações gráficas fez com que trouxesse este assunto para a aula. Mesmo na entrevista final, enquanto discutia a questão da localização geométrica da média, moda e mediana num gráfico de barras de uma distribuição enviesada, Estela revelou que adquiriu mais alguma capacidade para as reconhecer graficamente precisamente depois de experienciar, na sua prática, as resoluções das tarefas t10 e t15. Porém, o facto de na aula ter feito alguns raciocínios errados, quanto à localização da mediana e quanto à variabilidade dos dados na tarefa t15, fez com que não detetasse o erro no enunciado desta tarefa. Esta situação parece evidenciar a pouca experiência de Estela em analisar e em relacionar representações.

CC. No ensino deste tópico, Estela enfatizou os cálculos associados às noções trabalhadas e a determinação de representações numéricas e gráficas, sobretudo através da calculadora gráfica. A experiência de ter introduzido novas tarefas na aula terá contribuído para que ganhasse uma perspetiva diferente acerca das medidas de localização no currículo e, simultaneamente, alguma noção do desafio em lidar com essas noções graficamente, quer para si quer para o aluno.

5.3.4 Medidas de dispersão

Mathematics is often taught in school as being exact and precise
Statistics is about 'noise', that is, how to measure and control variability.
Burrill e Biehler (2011, p. 64)

A análise que se faz do conhecimento didático em Estatística de Estela, relativa a este tópico, tal como refletido na sua prática, divide-se em três secções. Na primeira (desvio-padrão), apresenta-se uma contextualização das medidas de dispersão e analisa-se como é que essas medidas foram desenvolvidas na prática de Estela; na segunda secção (reflexões sobre o tópico na aula), apresenta-se algumas reflexões da professora sobre o ensino e aprendizagem do tópico. Na terceira (discussão do tópico), mostra-se uma discussão do tópico guiada pelas quatro dimensões do modelo do conhecimento do professor em Estatística.

5.3.4.1 Desvio-padrão

Das medidas de dispersão, amplitude total, amplitude interquartil, variância e desvio-padrão, a que teve maior expressão no tópico foi, indubitavelmente, o desvio-padrão. A amplitude total e a amplitude interquartil foram mencionadas algumas vezes no ensino do tópico das medidas de localização (tópico anterior), nomeadamente, nas tarefas que pediam a determinação de quartis, construção do diagrama de extremos e quartis e algumas leituras desses diagramas.

Os alunos habituaram-se a ver o símbolo do desvio-padrão (s ou σ) que foi lido pela professora como desvio-padrão amostral ou populacional, respetivamente, através do seu trabalho na calculadora gráfica para determinar as medidas de localização, mesmo antes de aprenderem este conceito na sala de aula. O conceito de desvio-padrão foi introduzido pela professora, numa das primeiras aulas observadas, aquando da realização da tarefa *Temperaturas máximas de Lisboa e de Madrid* (t8).

Tarefa *Temperaturas máximas de Lisboa e de Madrid* (t8)

A professora introduziu e desenvolveu o conceito de desvio-padrão quando os alunos já se encontravam a trabalhar nesta tarefa e fê-lo assim que eles chegaram aos resultados da média, moda e mediana tal como solicitado na primeira alínea. Estela começou por indicar a página do manual onde se encontrava a fórmula da variância e mencionou que o desvio-padrão "tratava-se da raiz quadrada positiva da variância" (observação 3). Além disso, reescreveu a expressão no quadro embora de uma maneira menos condensada em relação àquela que constava no manual, sem o símbolo de somatório (ver figura 30).

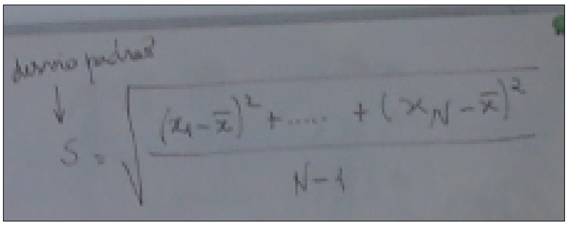


Figura 30: Fórmula do desvio-padrão (no quadro).

Também apresentou o desvio-padrão como uma medida de dispersão, "o desvio-padrão não é uma medida de localização, mas de dispersão, mede quanto é que os dados se distanciam, afastam, da média" (observação 3) e fez no quadro o cálculo desta medida com os dados das temperaturas máximas de Lisboa para exemplificar o uso da sua fórmula (ver figura 31):

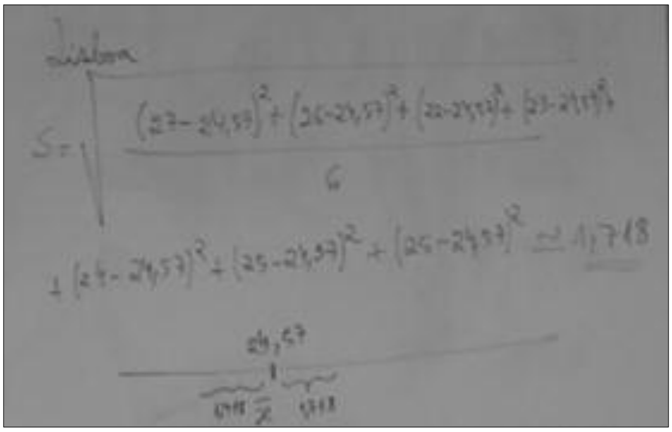


Figura 31: Cálculo do desvio-padrão com os dados de Lisboa.

Na interação com os alunos, Estela referiu ainda que eles deveriam determinar sempre o valor do desvio-padrão através da calculadora gráfica, uma vez que o cálculo à mão era moroso e não era exigido pelo programa do ensino secundário. Aproveitou também para lhes relembrar como se encontrava o valor do desvio-padrão na calculadora gráfica. Estela elucidou que esta medida estatística surgia juntamente com as outras estatísticas usuais (por exemplo, média, moda, mediana e quartis) quando se utilizava a mesma instrução de três passos (*STAT, Menu CALC, opção 1: 1-Var-Stat*) na calculadora gráfica para calcular, por exemplo, a média de um conjunto finito de dados. Adicionalmente, Estela esclareceu que, no caso do desvio-padrão, o seu valor encontrava-se habitualmente representado por *s* (desvio-padrão amostral), por se lidar, na maioria das vezes, com dados provenientes de amostras. Recomendou ainda que, caso tivessem de calcular o desvio-padrão (populacional) com dados de uma população, o correto seria usar denotar o resultado com *sigma* em vez de *s*. Estela não explanou, aos alunos, o motivo dessa recomendação.

A interação que a professora teve com os alunos prosseguiu à volta da interpretação do valor do desvio-padrão e sua aplicação com os dados de Lisboa:

Professora: A média para Lisboa era quanto? 24,57, certo?... para Lisboa…O que é que quer dizer um desvio-padrão de 1,718?
Mariana: Acho que é… as temperaturas de Lisboa estão a 1,718 de distância da média…
Professora: Pois… temos a média…e o valor do desvio-padrão diz que a distância entre cada um dos valores das temperaturas para a média é de 1,718… Ou seja, os dados estão mais ou menos a uma distância de 1,718 da média… da média para ali ou para aqui (regista, no quadro, num segmento o valor da média e a distância de 1, 718 à sua direita e à sua esquerda). Está?... Por acaso, não tiveram curiosidade de calcular os extremos deste intervalo de confiança média mais ou menos desvio-padrão, pois não?
Mário: Estou a fazer… Dá 22,85 Stora… o primeiro… e o segundo dá 26,29.
Professora: Temos aqui os dados… como podem ver a maior parte dos dados estão no intervalo, não é?... Até vos posso dizer mais… um dia vão aprender isto na universidade… Cerca de 68% dos dados estão aqui (apontando para o intervalo) … Não estão todos, pois não? O 27 está? Não… O 26 está!... O 22 não está!
Mário: Stora, então, o 27 é o desvio-padrão?
Professora: Não, não… o desvio-padrão é de 1, 718 como tínhamos calculado! O 23 está lá… O 25 e o 24 também! Então… ao todo, quantos estão [no intervalo]? São 5 ou 6?
Alguns alunos: São 5.
Professora: Cinco em quantos?… em sete… Cinco sétimos vezes cem dá em percentagem? À volta de 71%... No intervalo estou a considerar que as temperaturas de Lisboa… estamos a assumir que estão a seguir uma determinada forma… determinado modelo que permite concluir isso… Mas aqui não há garantias… Na universidade irão ver isso, está?… Então onde é maior o desvio-padrão [Lisboa ou Madrid]? Já calcularam?
Alguns alunos: É em Madrid, Stora…
Professora: E a média de Madrid?
Alguns alunos: É menor...
Professora: É menor. Pronto. (observação 3)

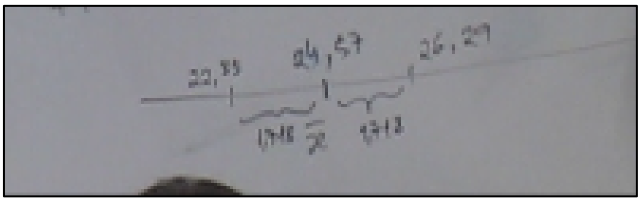


Figura 32: Intervalo de confiança registado no quadro.

Assim que indicou a fórmula do desvio-padrão, Estela explicou, de um modo informal e também superficial, o que entendia pelo desvio-padrão, descrevendo-o como uma medida de dispersão que mede, de uma maneira geral, o distanciamento dos dados à média. Nesta interação, talvez por influência da resposta da aluna Mariana, que indicou o valor desvio-padrão como uma distância entre dados e média, Estela usou os valores obtidos de média e desvio-padrão (do caso Lisboa) para construir um certo intervalo de confiança média mais ou menos desvio-padrão (ver figura 32) e para mostrar, através dele, que cerca de 68% dos dados da distribuição estariam inseridos no dito intervalo de confiança[24]. Quando fez essa verificação juntamente com os alunos, chegaram à conclusão que aproximadamente 71% dos dados estavam no intervalo. Contudo, alertou que esta distribuição teria de obedecer a um modelo específico para ser válida. Estela optou por omitir falar no modelo *normal* ou *gaussiano*, mas deixou a indicação de que este modelo era um assunto a ser estudado na universidade. Também não indicou à aluna Mariana que a sua resposta não era adequada (pois o desvio-padrão não é uma distância entre dados e média, mas será aproximadamente uma média destas distâncias) mas exemplificou o uso do desvio-padrão na constituição de um intervalo em torno da média que incluiria pelo menos 68% dos dados.

Estela questionou os alunos sobre qual das distribuições de dados (de Lisboa ou de Madrid) teria maior desvio-padrão. Alguns alunos responderam prontamente que o valor de Madrid era superior ao de Lisboa. No entanto, e apesar de o enunciado pedir explicitamente a interpretação dos resultados obtidos, não houve qualquer comentário da professora sobre o que teria influenciado a diferença de valores entre os desvios-padrão de Lisboa e de Madrid, nem uma interpretação dos resultados obtidos à luz dos dados e seu contexto.

Tarefa *Propriedades da média e do desvio-padrão* (t13)

Esta tarefa foi produzida pela professora. A sua resolução foi bastante orientada por ela, no entanto, Estela foi sempre avançando em interação com os alunos. Depois de ter lido o enunciado desta tarefa, a professora solicitou aos alunos que a resolvessem na calculadora. E foi ainda mais específica dizendo: "Eu vou-vos ensinar a calcular isso tudo (referindo-se à média, moda, quartis e desvio-padrão, tal como pedido na alínea a) de uma vez na calculadora gráfica!" (observação 5). De seguida passou a relembrar-lhes que deviam colocar os dois conjuntos de dados, salários e respetiva frequência, em duas listas diferentes associando, no quadro, a lista 1 aos salários (600; 700; 800; 1000; 1500; 2000; 4000) e a lista 2 às respetivas frequências absolutas (20; 50; 15; 9; 3; 2; 1).

[24] Pelo Teorema de Tchebyssheff, pode-se garantir que pelo menos 75% dos dados estarão inseridos no intervalo de confiança $]\bar{x} - 2s, \bar{x} + 2s[$, para qualquer conjunto amostral ou populacional de dados.

Projetou a calculadora gráfica na tela e mostrou como colocar os dados nas listas enquanto perguntava aos alunos as instruções precisas a dar à calculadora.

Uma das razões que pode ter levado a professora a centrar mais em si a resolução desta tarefa, não dando tempo aos alunos para trabalharem de modo autónomo, como procede habitualmente, foi ter-se apercebido que os alunos ainda não estavam suficientemente familiarizados com a calculadora gráfica. Além disso, eles não tinham ainda experiência em trabalhar com mais do que uma lista de dados na calculadora, o que, para o caso, era um conhecimento essencial, uma vez que cada dado do enunciado vinha acompanhado da respetiva frequência.

Observou-se, na interação entre professora e a turma, algumas dificuldades dos alunos na resolução desta tarefa, nomeadamente, na compreensão de que o acréscimo de 5% nos salários originais (alínea b) correspondia a multiplicar os valores destes salários por 1.05. Face a estas dificuldades, a professora avançou uma sugestão de desbloqueio ainda que vários alunos continuassem a achar que bastava multiplicar por 5%. Outra dificuldade que se notou foi na gestão das listas nas calculadoras. Alguns alunos tinham, na sua calculadora, dados de outras tarefas nas listas que a professora estava a sugerir que usassem e outros tinham algumas listas desaparecidas, ou com o conjunto de listas desordenado. Esta situação fez com que vários alunos não conseguissem obter as mesmas estatísticas que a professora estava a obter para cada situação indicada na tarefa. Por conseguinte, Estela despendeu algum tempo a esclarecer todas essas situações. Por exemplo, pediu que anotassem muito bem as listas de dados que estavam a usar para cada situação porque não era obrigatório que fossem as mesmas que ela usava. Estela acabou por realizar todos os cálculos na calculadora gráfica em simultâneo com os alunos (ver figura 33), usando o projetor com a sua calculadora, e os alunos, nos seus lugares, trabalharam a pares, tendo o cuidado de explicitar repetidamente os vários passos a seguir na calculadora.

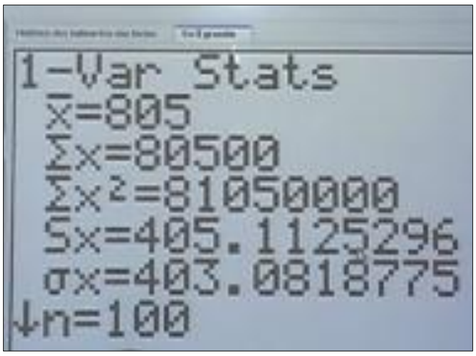


Figura 33: Ecrã da calculadora com algumas estatísticas dos dados originais.

Depois de todos os cálculos efetuados (média, moda, mediana, quartis e desvio-padrão para cada situação original e das outras duas situações descritas no enunciado) e de ter registado, no quadro, todas as estatísticas das três situações (dados originais, dados com um aumento de 5% e dados com um aumento de 50 euros), tentou que os alunos olhassem para os resultados e os comparassem com as estatísticas iniciais (figura 34).

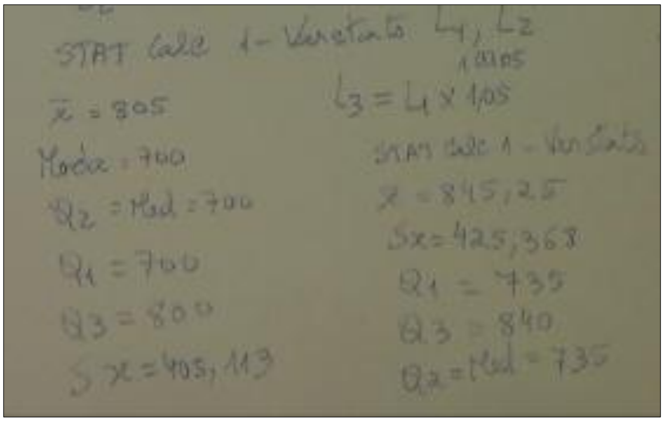


Figura 34: Colunas com estatísticas dos dados originais e dos dados que tiveram um aumento de 5%, respetivamente.

Como não houve alunos a avançar com respostas, Estela fez com que os alunos se apercebessem da influência de um acréscimo de uma mesma percentagem em todos os dados (ou um acréscimo de uma determinada constante nos dados) e o que é que isso provocava nas medidas estatísticas iniciais, concluindo a resolução da tarefa da seguinte forma:

> Então, moral da história: (…) se eu multiplicar os valores da variável estatística [dados] por um número, as medidas estatísticas vêm todas aumentadas por este número, aumentadas…, quer dizer… multiplicadas. Mas se adicionar aos valores

> da variável estatística [dados] um número, os valores das medidas estatísticas vêm todos adicionados deste número, exceto o desvio-padrão que se mantém. Está compreendido? Está mesmo? (observação 5)

Nem a professora nem nenhum aluno questionou porque é que é que o comportamento do desvio-padrão, quando os salários sofreram um acréscimo de 50 euros, se manteve inalterado, ao invés do das outras estatísticas cujos valores aumentaram 50 euros relativamente às estatísticas dos dados iniciais. Mesmo na outra situação em que os dados iniciais sofreram um aumento de 5% e os resultados obtidos aumentaram 5% comparativamente às estatísticas iniciais, tais comportamentos não foram questionados nem explicados no contexto de cada situação. A tarefa em si acabou por não se restringir a algumas propriedades da média e do desvio-padrão, uma vez que também foram comparados os valores da moda e dos quartis. Estela também sugeriu para trabalho de casa uma outra tarefa idêntica (*Alteração dos valores de uma distribuição*, t14) que incluía, além das medidas de localização usuais, as medidas de variabilidade: amplitude total e amplitude interquartil; contudo, a sua correção não foi feita na aula.

Tarefa *Distribuições das idades das equipas de trabalho* (t10)

A professora introduziu esta tarefa mencionando, de modo rápido, o que representa cada um dos diagramas de barras (ver figura 35), pediu também a colaboração dos alunos na identificação da notação de média e de desvio-padrão incluída no enunciado. Através desta tarefa, foi explorado o conceito de desvio-padrão em gráficos de barras. Professora e alunos tiveram a seguinte interação durante a sua correção:

> **Professora:** Já descobriram qual é a média?
> **Alguns alunos:** É 16 [referindo-se à média do gráfico A]
> **Professora:** Para aqui a média é 16. Onde é que também é 16?
> **Alguns alunos:** Na D.
> **Professora:** Então quer dizer que para o A ou para o D temos ou o [a hipótese I] 1 ou o [hipótese III] 3. Então agora basta ver qual deles tem maior desvio-padrão? É o A ou o D? (…) O Mário diz que que é o A [gráfico]. Tem maior desvio-padrão. Não se esqueçam da fórmula.
> **Jaime:** É o D [gráfico], stora. É o D! [que tem maior desvio-padrão].
> **Professora:** Pensem. Podem experimentar, estou a ver alguns a experimentar com a calculadora… Claro que eu queria que tivessem sensibilidade, que olhando… tentassem aventar uma hipótese para ver quem tem maior desvio-padrão. Mas podem experimentar. Aqui [referindo-se ao gráfico A] têm todos a mesma frequência. Vocês é que escolhem o valor [da frequência] e ali [referindo-se às barras do gráfico D] há uns que tem uma grande frequência, outros dois tem a

> mesma… mas é para aí metade ou menos… e as outras barras tem uma frequência ainda menor. (observação 5)

Nesta interação, a professora indica as associações que pretende que os alunos realizem, ou seja, que liguem cada gráfico à respetiva média e desvio-padrão (representados no enunciado da tarefa por um par ordenado numérico). Os alunos chegaram às respostas corretas para a média, localizaram os dois gráficos que tinham como média 16 anos e os restantes dois de média 16,3 anos. Contudo, não surgiu qualquer interação que evidenciasse o raciocínio que os alunos fizeram para chegar a essas conclusões.

A professora também sugeriu uma estratégia específica para os alunos determinarem o desvio-padrão que dizia respeito aos dados representados no diagrama A: que tomassem como altura das barras um valor qualquer para o diagrama A (o diagrama A fornecido não possuía essa informação, assim como os restantes diagramas – ver figura 35) e que determinassem, através da calculadora gráfica, o valor do desvio-padrão. Depois de propor que a utilizassem na determinação do desvio-padrão para os dados representados no gráfico de barras D, há a seguinte interação na aula com a turma:

> **Mário:** Stora, eu acho que a [hipótese: 1.05] I é a D e acho que a [hipótese: 1.41] A é o III.
> **Professora:** Onde tem maior [valor de desvio-padrão]? Achas que o A é o III. O Pedro já experimentou, não foi?
> **Pedro:** Sim.
> **Professora:** O Pedro experimentou… na calculadora… o caso em que todos tinham a mesma frequência e viu que dava aquele. (…) Vamos pensar porquê, que é para ver se conseguimos o B e o C agora direitinho, sem andarmos a experimentar. Lembram-se da fórmula? (Estela escreve a fórmula do desvio-padrão no quadro) Reparem. A média é 16. Aqui há muitos catorzes (apontando para a barra dos 14 anos do gráfico A)… portanto a distância deste ao 16 é maior [no gráfico A] do que por exemplo do 15 ao 16 [no gráfico D]. É ou não é?... É. Se tiver o desvio-padrão, ele mede o quê? A distância de cada um dos valores da variável estatística em relação à média. Aqui como são muitos [apontando para as barras do diagrama A e indo assinalando com a mão a distância de cada barra à barra dos 16 anos], as diferenças são muito grandes, enquanto que aqui [no gráfico D], por exemplo, neste (apontando para as barras dos 15 e 16 anos no gráfico D) são menos ou são mais? Então quem tem mais desvio-padrão?
> **Ivo:** é o A.
> **Professora:** É este [A] ou o último [D]?
> **Alunos:** É o [gráfico] A.
> **Professora:** Como o Mário tinha dito, logo desde o início, sem experimentar. Foi com esta noção, não foi? (ouve-se vários risos, a professora assinala no quadro os desvios-padrão sobre os gráficos A e D, ver figura 20)… Então qual tem maior desvio-padrão (apontando para os gráficos B e C)?

Alunos: É o B. (observação 5)

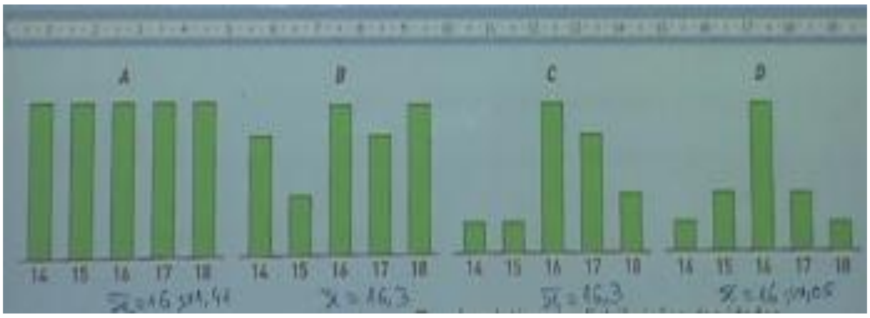


Figura 35: Os quatro gráficos (projetados no quadro) com os desvios-padrão assinalados nos gráficos A e D.

A estratégia referida acima para se chegar ao valor do desvio-padrão acabou por ser discutida superficialmente. Quando Estela sugeriu que o aluno Pedro teria usado essa ideia, crê-se que esta situação consistiu numa forma de chamada de atenção para o trabalho que tinham em mãos. Os valores de desvio-padrão que os alunos obtiveram efetivamente na calculadora gráfica não emergiram na interação entre professora e alunos, ficando subentendido que realmente foram obtidos os resultados que estavam no enunciado, ou seja, um valor de desvio-padrão para o grupo A de 1.41 e para o grupo B de 1.05. No entanto, a estratégia acima avançada pela professora parece ter permitido que os alunos calculassem e comparassem os desvios relativamente à média, e que identificassem qual dos gráficos A ou D tinha maior desvio-padrão. Em particular, desta forma, os alunos também poderiam confirmar o valor da média dos gráficos A e D, embora não tenha ocorrido nenhuma observação nesse sentido.

Na sua interação com os alunos, Estela também parece tentar fazer com que os alunos se apercebam que cada distribuição representada no respetivo gráfico de barras corresponde a amostras de tamanhos diferentes. Além disso, com o intuito de os ajudar a entender qual das distribuições, A ou D, tem maior desvio-padrão (tal como se mostra na interação acima), Estela referiu-se às “diferenças”, ou seja, aos desvios de cada dado (idade) em relação à média quando analisava esses gráficos. Contudo, os valores concretos dessas diferenças ou desvios à média (0 anos, 1 ano, ou 2 anos) não emergiram na discussão, o que poderia facilitar a compreensão do seu discurso dado que, por vezes, usa a expressão “as diferenças são muito grandes” não para assinalar a variabilidade dos valores das diferenças mas para se referir a uma maior frequência de determinados valores de diferença (ou desvios à média). Também, a partir do momento que os alunos forneceram as respostas corretas, identificando o gráfico com o respetivo valor de desvio-padrão, a professora não os questionou sobre o porquê das suas opções, talvez por considerar que já tinha avançado com as explicações necessárias que estes precisavam para entender.

5.3.4.2 Reflexões sobre o tópico na aula

Conceitos

O conceito de desvio-padrão é, para Estela, um conceito que tem vindo a desenvolver mais recentemente na sua prática profissional e sobre o qual tem também falado com alguns colegas de Matemática. Por exemplo, na entrevista inicial, Estela comentou que o conhecimento que, através da calculadora gráfica, é possível obter dois resultados para o desvio-padrão (desvio-padrão populacional e o amostral) e o modo como esses dois valores deveriam ser aplicados na resolução das tarefas é algo que os professores eventualmente conhecem há pouco tempo:

> Porque às vezes surge na máquina gráfica… quando estou a calcular o desvio-padrão aparece *Sx* e o σx. A Sandra lembra-se que há 2 ou 3 anos, num exame de MACS, houve uma grande questão porque os alunos escolhiam o valor do sigma [desvio-padrão populacional] mas, por que no enunciado dizia que era uma amostra, então [os alunos] tinham era de considerar o *Sx* [desvio-padrão amostral]. E quem eram os professores que estavam atentos a isso? Não havia. (…) Sabe onde ouvi isso? No mestrado, nunca na faculdade [na formação inicial]! Está bem que os estudos vão avançando e vão melhorando. (…) E agora chega-se [a essa conclusão]. Os manuais [também] não faziam isso. (entrevista 1)

Na entrevista final, após o término de observação de aulas, Estela referiu que, quanto ao conceito de desvio-padrão, tinha a perceção que os seus atuais alunos de Matemática A adquiriram o conhecimento que, dados dois conjuntos de dados numéricos, "aquele que tiver maior desvio-padrão, para a mesma média, é aquele em que os dados estão mais dispersos em relação à média. É o que eu entendo… Acho que os alunos ficaram com essa ideia intuitiva" (entrevista 3). Assim sendo, Estela tem a convicção que os alunos entendem o desvio-padrão como uma medida que fornece indicações quanto à maior ou menor dispersão dos dados relativamente à média.

A investigadora mostrou a Estela a situação *Desvio-padrão elevado*[25] (anexo 2) para ela a comentar. Ao ler a questão, a professora pergunta à investigadora se há algum gráfico a acompanhar a pergunta, ao que a investigadora responde negativamente. Estela comenta cada afirmação proposta na situação do seguinte modo:

> [Lê a primeira afirmação, e diz]… É verdade! [na segunda, diz] Aqui não está a ser usada só uma variável, são duas... [na terceira, diz] Está bem!… [na quarta, diz]… Ah! Que engraçado!... Realmente basta um *outlier* para aumentar o valor do desvio-padrão… altera logo… Lembro-me que até saiu um exercício em MACS num exame [dizia algo do género]: “Considera que um valor é um *outlier* se estiver… estiver fora do intervalo [com extremos] tal e tal. Calcula esse intervalo de confiança e diga se o valor tal… resultado… é um *outlier*. Achei engraçado!... Quanto às afirmações, só tenho dúvidas nesta (referindo-se à segunda). Acho que este argumento não é muito correto (…) Aliás… num contexto de dados bivariados calcula-se o desvio-padrão de cada variável, não é um desvio-padrão para as duas. (entrevista 3)

Estela denota compreender a influência de um *outlier* no cálculo do desvio-padrão e possuir ideias informais sobre este conceito.

A opção de lecionar o tema Estatística no 10.º ano recorrendo mais às representações gráficas foi uma decisão que tomou influenciada pela sua experiencia na ação de formação sobre Estatística em que participou. De todo o trabalho que desenvolveu nesta formação, a parte de lidar com as representações foi a que julgou mais pertinente e de utilidade imediata para a sua prática: “Neste ano letivo, na Estatística é novo explorar mais as representações, que é aquilo que nos salta mais à vista no dia-a-dia (…) eu não tive nada disso [enquanto aluna]” (entrevista 2). Analisar o desvio-padrão num gráfico de barras ou histograma era uma atividade que ainda recentemente considerava difícil e, por isso, não a propunha na sala de aula. O ter iniciado este tipo de análise com as medidas de localização média e mediana, na formação, facilitou que se desse a explorar

[25] *Situação Desvio-padrão elevado* – Englobava a questão, “O que poderia significar um desvio-padrão elevado para um conjunto de dados?” e as suas quatro respostas hipotéticas: 1. Os dados estão bastante dispersos entre si; 2. Um diagrama de pontos com os pontos a variar bastante entre si; 3. Os dados estão bastante afastados da média; e 4. Pequena variabilidade nos dados com a existência de um ou mais *outliers* (anexo 2).

o desvio-padrão em gráficos de barras. Daí ter selecionado, de um manual diferente do adotado, a tarefa *Distribuições das idades das equipas de trabalho* (t10) e tê-la proposto numa aula**.**

Na opinião de Estela, a Tarefa *Propriedades da média e do desvio-padrão (t13)*, que preparou para ser proposta na sala de aula, acabou por ficar muito bem conseguida por permitir que se fizesse um estudo à média e ao desvio-padrão a partir de uma mesma situação problemática, ao contrário das situações que encontrou no manual escolar adotado. Na sua perspetiva, este manual apenas incluía situações independentes para o estudo de propriedades da média e para o estudo de propriedades do desvio-padrão. Estela considerava útil que essas propriedades fossem analisadas num mesmo contexto ou situação problemática, crendo que isso poderia ajudar os alunos a alcançar uma perceção mais global de cada propriedade. Além disso, também considerava que, com esta tarefa, os alunos poderiam determinar os resultados da média e do desvio-padrão de uma só vez através na calculadora gráfica e, a partir daí, eles teriam mais tempo para retirar conclusões sobre os resultados, tal como disse: "… com estes resultados da calculadora poderiam averiguar o que é que acontecia com a média ou com o desvio-padrão quando se somam ou multiplicam os dados por um determinado valor, do que ser eu a dizer [as conclusões]" (entrevista 2). Estela tinha a expetativa que esta tarefa, na sua primeira parte, iria permitir que os alunos estudassem duas propriedades da média e duas propriedades do desvio-padrão dentro de uma situação específica. A segunda parte contribuiria para que a generalização dessas propriedades fosse pensada e refletida pelos alunos.

Produções dos alunos

Como já foi escrito anteriormente, a tarefa *Relatório sobre os resultados do teste intermédio (t12)* acabou por ficar para trabalho de casa e a professora recolheu os relatórios dos alunos que elaboraram o relatório pedido. Dos 25 alunos, apenas 11 entregaram este trabalho de casa. Destes relatórios, há apenas três em que foi incluído o resultado do desvio-padrão. Mais concretamente, destes três, há dois em que indicaram um resultado errado (de aproximadamente de seis valores) e um com o resultado certo (de 3,28 valores). Destes três alunos apenas um justificou a sua resposta, a aluna Cátia, que a professora descreve como uma aluna "muito trabalhadora" (entrevista 2), cuja resposta foi:

> [Desvio-padrão] (s), s=6, 352, aproximadamente 6 (segundo a máquina gráfica). Quanto maior for o desvio-absoluto, mais afastado se encontrará do zero, consequentemente maiores serão os quadrados dos desvios que o origina uma maior dispersão dos dados em relação à média. Neste caso, os resultados obtidos não são muito dispersos visto se encontrarem perto da média obtida.

Enquanto a professora analisa essa resposta da aluna, tem a seguinte interação com a investigadora:

> **Professora:** Ela calculou o desvio-padrão na máquina [calculadora] gráfica… Quanto maior for o desvio absoluto?... Mais afastados do zero? O que é isto? Ah! (Lê novamente a resposta)… Aqui deve estar escrito visto (visto não estar muito legível)… mas não foi a linguagem que se usou na aula!... Não se falou no desvio absoluto nem nos quadrados dos desvios relativamente ao desvio-padrão.
> **Investigadora:** Ela deve ter visto nalgum manual a definição de desvio-absoluto médio …
> **Professora:** Pois, ela [Cátia] procura muito… (entrevista 2)

A professora acaba por não achar adequada a resposta da aluna Cátia, sobretudo por considerar que não está em conformidade com o modo como o desvio-padrão foi expresso na aula. Estela disse, por exemplo, que não foi abordada a definição do desvio médio absoluto. Contudo, este conceito emergiu de forma implícita no desenvolvimento da tarefa *Distribuição das idades das equipas de trabalho (t10),* quando a professora explicava aos alunos os raciocínios que poderiam fazer para determinar, entre duas distribuições representadas por gráficos de barras, qual a de maior desvio-padrão. Estela, por sua vez, também acaba por revelar estar bastante dececionada com o relatório de Cátia visto ter encontrado vários erros[26] que lhe pareceram ser oriundos de uma interpretação errónea da tabela dos dados, fornecida no enunciado.

[26] Estela detetou que esta aluna apresentou um diagrama de extremos e quartis com os valores dos quartis errados, um diagrama de caule e folhas em que não se entendia os dados e os valores que usou e ainda média e mediana mal calculadas. A aluna Cátia supostamente calculou o valor do desvio-padrão (embora não tendo obtido o valor correto que é de cerca de 3 valores) na calculadora, referindo, no seu comentário, o desvio absoluto (médio); porém, não o calculou. Também não se sabe se escreveu por lapso desvio absoluto em vez de desvio-padrão. Acrescenta-se ainda que, com esta resposta que deu, a aluna parece querer estabelecer alguma relação entre desvios, quadrados dos desvios e o desvio-padrão mas, quando se refere aos "resultados obtidos", não clarifica quais são esses resultados (será que se refere aos desvios? Ou será aos quadrados dos desvios?, etc.). Por conseguinte, com o seu comentário não se entende como é que ela avalia a dispersão dos dados.

A tarefa *Aproveitamento das turmas X e Y*[27]*(t11)*, que foi selecionada do manual e sugerida para trabalho de casa, inclui dois gráficos de barras com a distribuição das notas da turma X e da Y. Ambos os gráficos são simétricos: enquanto no gráfico da turma X, a altura das barras decresce das extremidades para o centro, no gráfico da turma Y, a altura das barras cresce das extremidades para o centro. É de referir que dezassete alunos (68% dos alunos da turma) entregaram este trabalho de casa. Estela pensa que os alunos não teriam grandes dificuldades em responder às diferentes alíneas desta tarefa por reconhecer as suas semelhanças com a tarefa *Distribuições das idades das equipas de trabalho* (anexo 10) que foi trabalhada na aula. Antecipa algumas dificuldades que poderiam surgir nos alunos relativamente à leitura das alturas das barras "quem me diz a mim que a barra está mesmo no 1 e não mais abaixo?" (entrevista 2), que era algo que os alunos tinham de verificar e assinalar para chegarem aos resultados de quartis e desvio-padrão pedidos na alínea c). Ao analisar todas as respostas dos alunos, a professora ficou satisfeita com o facto da maioria dos alunos ter chegado aos resultados certos através da calculadora gráfica. Contudo, estava à espera que a resposta final (alínea d) fosse mais completa e assertiva por parte dos alunos e que envolvesse mais interpretação dos resultados, concluindo que: "daqui vejo que…

[27] Da análise sucinta às respostas apresentadas por estes alunos destaca-se o seguinte:

- Todos os alunos identificaram que a média das duas turmas era 13 (alínea a), dos quais cinco alunos apresentaram o algoritmo da média para cada turma, dois alunos associaram a média 13 com a classificação das barras centrais dos dois gráficos barras, e os restantes alunos indicaram apenas que a média das duas turmas era 13, sem apresentar argumentos justificativos.
- Na previsão de qual seria a turma com maior desvio-padrão (alínea b) todos indicaram que a turma X deveria ser a que tinha maior desvio-padrão devido a ter mais classificações afastadas da média.
- Na alínea c, no cálculo do desvio-padrão, mediana e quartis, quatro alunos apresentaram valores errados de desvio-padrão e quartis e os restantes apresentaram os valores corretos. Todos assinalaram que usaram a calculadora gráfica para obter esses resultados.
- Na alínea d, na comparação do aproveitamento das duas turmas, há apenas um aluno que não responde e uma outra aluna que diz "que o aproveitamento de ambas é equivalente" (aluna Laura). Os restantes (15 alunos) apontam que a turma com melhor aproveitamento é a Y, embora as justificações variem. Destes alunos, há alguns que indicaram que a igualdade das médias das duas turmas os obrigou a procurar outro critério de desempate, optando pelo desvio-padrão. Há também alguns alunos que descrevem um pouco mais a distribuição das notas da turma Y e destes a resposta mais frequente engloba esta ideia transmitida pela Sara: "Há menos alunos maus, mas também há menos alunos bons, no entanto uma boa parte dos alunos é razoável a bom"; outros exemplos são o caso do aluno Jaime que escreveu "(…) 75% da turma Y tem notas entre 11,5 [1.º quartil] e 17 [máximo] enquanto na outra turma [X, os mesmos valores] se situam entre 10 e 17" e ainda a resposta do aluno Mário: "A turma Y tem menos negativas". Há também duas alunas que assumem a turma Y como a melhor mas, nas suas respostas, destacam antes aspetos da turma X, nomeadamente ser a que tem maior frequência na melhor nota (que é 17): "…no entanto, a turma X tem maior número de bons alunos" (resposta da Mariana).

oralmente… na aula… discutimos todos os assuntos tratados, mas por escrito era preciso trabalhar mais um bocado a escrita das conclusões e suas explicações!...” (entrevista 2).

5.3.4.3 Discussão do tópico

CETE. Estela propôs tarefas diversificadas que acabaram por ser desenvolvidas em interação com os alunos mas com muita orientação da sua parte. Na sua atuação, houve mais ênfase nos cálculos e na utilização da calculadora gráfica em detrimento da explicação de ideias estatísticas envolvidas nas tarefas ou de se fazer um relacionamento entre os resultados para melhor entendimento dos dados e do seu contexto. Por exemplo, na tarefa *Temperaturas máximas de Lisboa e de Madrid* (t8), os valores dos desvios-padrão foram alcançados, mas a comparação desses valores tendo em conta o contexto dos dados, que era solicitada no enunciado, não foi explorada. Outro exemplo é o da tarefa *Propriedades da média e do desvio-padrão* (t13) que foi trabalhada com foco no cálculo de medidas estatísticas, envolvendo também a comparação desses resultados, obtidos na calculadora gráfica (assim que se foram alterando os dados iniciais). A análise realizada foi muito dirigida pela professora, embora os alunos, de uma maneira geral, estivessem a seguir e a concordar com o que estava a ser dito. Na interação promovida com vista ao entendimento de relações entre medidas calculadas, à medida que os dados originais iam sofrendo alterações específicas, houve sobretudo uma certa sobrevalorização dos resultados obtidos na calculadora gráfica, que poderá ter influenciado a falta de necessidade de os questionar, explicar e relacionar uns com os outros. Ainda assim, as duas propriedades da média e do desvio-padrão anunciadas no fim da tarefa parecem ser consideradas pela professora como relevantes para um entendimento mais complexo destas estatísticas. Mesmo na tarefa 13, que iria permitir estudar o desvio-padrão graficamente, a professora sugeriu que os alunos atribuíssem valores à altura das barras para determinarem, na calculadora gráfica, uma estimativa do valor do desvio-padrão. Neste caso, também abordou brevemente a variabilidade dos dados face à média para abordar o desvio-padrão; no entanto, por vezes, foi ambígua nas suas explicações. A sua participação na ação de formação terá contribuído para que selecionasse algumas tarefas para a aula que envolviam comparações entre representações de dados, que nunca tinha experimentado no contexto do ensino da Estatística. Esta situação tem repercussões relevantes ao nível do seu CETE.

CA. As dificuldades observadas dos alunos em lidar com a calculadora gráfica na exploração da tarefa *Propriedades da média e do desvio padrão* (t13), que foram logo inicialmente detetadas pela professora, podem dever-se ao facto de, naquela altura, ainda haver alguns alunos que estavam a usar pela primeira vez a calculadora gráfica e haver outros que ainda não tinham tido tempo suficiente para assimilar instruções que já tinham sido referidas algumas vezes nas apresentações dos grupos e pela professora na resolução e correção de tarefas anteriores. Estela foi detetando algumas dificuldades dos alunos em abordar as tarefas propostas. Muitas vezes, nessas alturas, foi sugerindo estratégias de resolução da situação. Tal foi o caso com a tarefa *Distribuições das idades das equipas de trabalho* (t10), em que recomendou, logo de início, que os alunos tentassem usar a calculadora gráfica para calcular o desvio-padrão das distribuições representadas em gráficos de barras (cujos valores das alturas das barras não estavam identificados), mesmo quando alguns alunos já estavam a associar valores de desvio-padrão com os gráficos, optando por não os interpelar acerca do modo como estavam a pensar. De uma maneira geral, na sala de aula, houve um acesso relativamente limitado aos raciocínios e explicações dos alunos. Quando a aluna Mariana lhe descreveu o desvio-padrão como uma distância, Estela não corrigiu a sua descrição inadequada. Quando Estela foi confrontada com respostas aos trabalhos de casa dos seus alunos, ficou surpreendida por encontrar erros nalgumas delas e pela maioria não incluir justificações, mesmo nas alíneas que o solicitavam. Estela considerou que o relatório da tarefa 12 exigia um comentário global que envolvesse todos os cálculos e representações efetuadas. Da análise destes relatórios, considerou que a parte de escrita de conclusões e comentários não estava tão bem conseguida, devido a este tipo de trabalho não ter sido muito desenvolvido na aula.

CE. Estela asseverou que ter ensinado MACS, ter frequentado um mestrado Educação Matemática, em que estudou alguns temas estatísticos, e ter feito várias reflexões em torno da Estatística, individualmente e com alguns colegas de Matemática, a terão ajudado a profundar os seus conhecimentos sobre medidas de dispersão, nomeadamente, sobre o desvio-padrão. Sobre este conceito, conhece a sua fórmula, sabe como o determinar na calculadora gráfica para o caso do conjunto de dados ser a população ou ser uma amostra, sabe construir um intervalo de confiança para a média amostral (usando o desvio-padrão), sabe que é uma medida pouco resistente a valores extremos (embora não tenha referido essa propriedade na aula); contudo, o traduzir este conceito por palavras suas foi um exercício que realizou efetivamente com a investigadora na entrevista inicial. A ideia informal de desvio-padrão como indicador da distância entre dado e média, que a professora indicou talvez por influência de uma aluna, não era adequada. Estela revela algumas dificuldades no entendimento da noção de desvio-padrão, pelo que, na aula, pode ter debilitado o conhecimento do aluno neste assunto.

CC. Neste tópico, Estela deu principal destaque ao desvio-padrão, relacionou-o com a média e abordou algumas das suas propriedades. Deu relevância à exploração da calculadora gráfica para se obter os resultados pretendidos nas tarefas. A sua participação na ação de formação terá contribuído para que sugerisse nas aulas tarefas que envolviam comparações entre representações, que prestassem atenção ao desvio-padrão.

5.3.5 Dados bivariados e regressão linear

*A palavra **regressão** deve-se a F. Galton que observou que:*
*Filhos de pais mais **altos**/baixos do que a média tendiam a ser mais **altos**/baixos do que a média, mas mais **baixos**/altos do que os pais, havendo assim uma tendência geral para "**regressar**" aos valores médios da população.*
(adaptado de Pestana &Velosa, 2010, p. 185)

Este tópico é analisado em quatro secções. As duas primeiras (correlação e regressão linear, e modelo linear) incidem sobre algumas noções centrais, explicam o modo como essas noções emergiram e foram desenvolvidas na prática de Estela. A terceira secção (reflexões sobre o tópico na aula), apresenta algumas reflexões de Estela acerca do ensino e aprendizagem deste tópico; a quarta secção (discussão do tópico), expõe uma discussão do tópico guiada pelas quatro dimensões do modelo do conhecimento didático do professor em Estatística.

5.3.5.1 Correlação e regressão linear

Nesta secção, foram definidas quatro partes (dados bivariados; reta de regressão; centro de gravidade; e coeficiente de correlação linear) que indiciam, em particular, a sequência e enquadramento das noções associadas à correlação e regressão linear tomada por Estela na sua prática. Na secção, também se dá destaque à tarefa *Equipa de basquetebol do Porto (t16)* visto ter sido preparada e proposta na aula especialmente para introdução de todos os conceitos associados ao tópico.

Dados bivariados

Tarefa *A equipa de basquetebol do Porto (t16)*

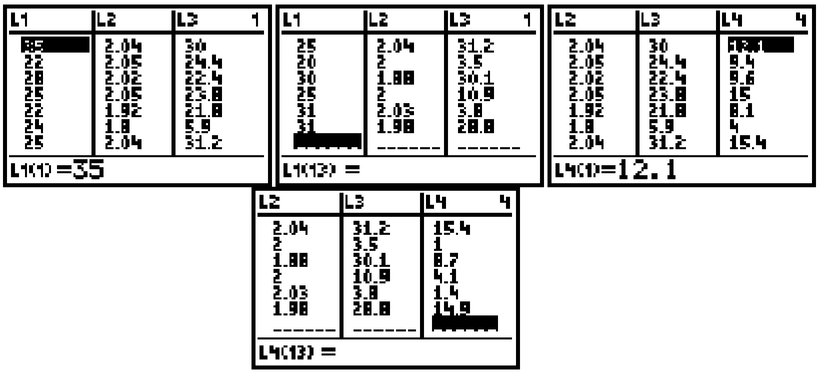
Ao ler cada uma das três questões principais do enunciado desta tarefa, Estela perguntou aos alunos se achavam ou não que elas "faziam sentido" no contexto em que se inseriam. A maioria dos alunos expressou que a primeira e a última pareciam referir-se a relações que poderiam ser analisadas, enquanto a relação exposta na segunda questão, ou seja, entre as variáveis *idade dos jogadores* e *pontos marcados* lhes parecia que não. A professora acabou por não dar seguimento a essas ideias transmitidas por alguns alunos, o que não permitiu perceber em que argumentos se baseavam, em particular, se tinham em conta a sua opinião pessoal sobre o assunto ou se faziam alguma análise aos dados.

Estela aproveitou ainda essas questões do enunciado para evidenciar que, em cada uma delas, estavam a ser relacionadas duas variáveis em simultâneo, em vez de se considerarem situações univariadas, como anteriormente. As variáveis bidimensionais foram apresentadas como variáveis cujos valores numéricos poderiam ser representados graficamente através de pontos:

> **Professora:** [Nas questões] Quanta informação se está a relacionar de cada vez?
> **Alguns alunos:** Duas.
> **Professora:** Duas… Quando faço isto, quer dizer… que estou a estudar variáveis bidimensionais… Ou seja, à mesma pessoa [jogador] eu estou a fazer uso de duas informações à custa dela. Cada conjunto de duas informações [de um mesmo jogador] corresponde graficamente a um ponto… em que uma das coordenadas é uma das variáveis que nós é que escolhemos e a outra coordenada é a outra informação. Certo? Tentem seguir essas instruções (referindo-se inicialmente ao diagrama de dispersão, usando os dados da terceira questão da tarefa) para depois me dizerem o que é que eu faço na máquina… (observação 5)

Estela pediu algumas vezes que, na exploração desta tarefa e numa primeira abordagem no que diz respeito aos dados a analisar, se focassem nos dados envolvidos na terceira questão do enunciado, ou seja, na questão que se refere "à eficácia do jogo", mais concretamente às variáveis: "minutos de jogo e pontos obtidos" (observação 5). Solicitou também que os alunos seguissem um conjunto de orientações, para utilização da calculadora gráfica, incluídas na tarefa, nomeadamente, para fazer a inserção de dados nas listas e obter o diagrama de dispersão (ver figura 36) e ainda a determinação de medidas estatísticas de cada uma das variáveis em simultâneo (cujo excerto é apresentado na secção seguinte).

A. Introdução dos dados bivariados nas listas (STAT + EDIT):



B. Diagrama de dispersão (2ND + STAT PLOT):

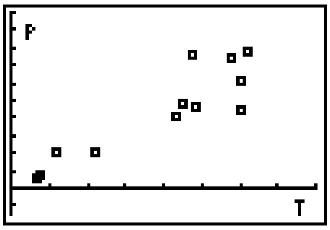


Figura 36: Excerto que mostra a inserção de cada variável na sua respetiva lista e o diagrama de dispersão.

Por motivos técnicos, a professora não conseguiu conectar a calculadora gráfica ao seu computador como desejava. Em alternativa, decidiu fazer o diagrama de dispersão no quadro com os marcadores, acabando por assinalar apenas alguns dados. Desenhou um referencial, colocando no eixo das abcissas os *minutos* e no eixo das ordenadas os *pontos* obtidos por cada jogador:

> **Professora:** Para o *Anthony Blackely* temos 30 minutos e 12.1 de pontos, certo? Temos de marcar aqui [no referencial] este ponto. (Depois de marcar este ponto) Este é o ponto que representa o senhor *Anthony*… O ponto que

> representa o senhor *Elvis Évora* é 24,4 que é para aí aqui … e o 9,4 aqui (desenha esse ponto). E assim sucessivamente…. Quantos pontinhos vamos ter [no total]?
> **Alguns alunos:** 12 pontos…
> **Professora:** A este gráfico que vão acabar de fazer chama-se diagrama de dispersão ou nuvem de pontos. … Vão fazer agora à mão, depois ensino na máquina gráfica. (Ao passar pelo aluno Leonardo apercebe-se que este aluno tinha realizado todas as instruções na calculadora *TI-Nspire* e diz a toda a turma) O Leonardo já fez tudo na calculadora!… Fez o diagrama de dispersão e até a reta de regressão! Já fez tudo! Já fizeram o gráfico? (observação 5)

O diagrama de dispersão foi apresentado como um gráfico que inclui os dados bivariados. Quando a professora pediu aos alunos para terminarem de assinalar manualmente todos os dados bivariados no gráfico, já alguns deles tinham conseguido obter o diagrama de dispersão na calculadora gráfica. Estela reparou, em particular, que o aluno Leonardo, naquele momento, tinha encontrado, por si mesmo, uma forma de representar o gráfico e a reta de regressão na calculadora mais avançada que possuía.

A professora fez, através do diagrama de dispersão, a seguinte leitura relativamente à associação entre as variáveis *tempo* e *pontos*:

> **Professora:** Já fizeram o gráfico [à mão]?
> **Alguns alunos:** Sim.
> **Professora:** Olharam para o que já fizeram?… E o que é que reparam?… Acontece que, à medida que o *tempo* aumenta, vocês veem que os *pontos* também aumentam ou não?
> **Alguns alunos:** Sim…
> **Isabel:** Há exceções!
> **Professora:** Então diz… Há exceções…mas a maioria (dos pontos)… Já fizeram na calculadora este diagrama?
> **Vários alunos:** Não. (observação 5)

Estela voltou a comentar o mesmo gráfico, depois de a maioria dos alunos ter conseguido colocar o diagrama de dispersão no visor das suas calculadoras gráficas:

> Então estão a ver que… parece que se consegue fazer passar uma reta não por todos os pontos… ora bem… mas pela maior parte deles ou, pelo menos, mais próximo deles. Então agora só falta descobrir como é que se desenha essa reta, certo? Essa reta chama-se reta de regressão, que é uma linha, uma reta, há vários tipos de regressão, linear, quadrática, cúbica, etc…, mas aqui estamos a tratar a linear… (Observação 5)

Através dos dois últimos excertos, observou-se que Estela começou por fazer uma leitura da relação bivariada expressa no diagrama de dispersão, procurando caraterizá-la de uma forma global: assim que os valores de uma variável aumentam (variável *tempo* - eixo *xx*) os valores da outra também aumentam (variável *pontos* - eixo *yy*). Associou essa tendência global à possibilidade de se fazer passar uma reta que se aproxime da maior parte dos pontos, designando-a por reta de regressão. Estela apontou, também, que há outros tipos de regressão (quadrática, cúbica) além da linear, que era, contudo, a única que iria ser estudada no secundário.

Reta de regressão

Tarefa *A equipa de basquetebol do Porto* (t16)

Estela pediu, ainda, a Leonardo, o único aluno que se voluntariou, para ir ao quadro desenhar o diagrama de dispersão dos dados bivariados associados às variáveis *tempo de jogo* e *pontos marcados* da terceira questão da tarefa. O aluno desenhou também, a pedido da professora, no mesmo gráfico, a reta que ele achava que visualmente era a "que melhor se ajustava aos pontos do gráfico", nas palavras de Estela (observação 6). Ao considerar que o Leonardo não a desenhou bem, Estela reagiu assim: "eu nunca traçaria esta [apontando para a reta desenhada sobre o diagrama de dispersão] porque … isto nem parece uma reta…são tantos [segmentos de] retas!" (observação 6). Por conseguinte, Estela optou por desenhá-la outra vez. Naquele momento, o Leonardo pediu à professora para fazê-la passar pelo ponto de coordenadas (22,4; 9,6), que se referia aos dados *minutos de jogo* e *pontos obtidos* referentes ao *jogador José Pedrera*, explicando considerar que este dado bivariado era indubitavelmente "um ponto da reta que procurava pelo esboço" que tinha feito no seu caderno diário (Observação 6). Outros dois alunos também expressaram a sua concordância com esta indicação do Leonardo, mas não lhes foi perguntado os seus motivos.

Estela desenhou a suposta reta que melhor se aproxima dos pontos a passar pelo ponto que os alunos referiram e por um outro que ela indicou: "o ponto de coordenadas: média dos valores da variável *minutos de jogo* e média da variável *pontos obtidos*" (observação 6), que foi de imediato designado por centro de gravidade, pelo aluno

Jaime. Ao pedir as coordenadas deste ponto, surgiram várias respostas diferentes. Por conseguinte, Estela alertou para que tomassem atenção às listas usadas, pois considerava que os erros que estavam a cometer provinham do facto de não estarem a usar as listas corretas na calculadora. Pediu depois para confirmarem novamente as listas que colocaram à frente da instrução que permite calcular a média de dois conjuntos de dados em simultâneo (*2- Var Stats*), explicitando que deveriam colocar primeiro a lista correspondente à variável *pontos obtidos* e de seguida a da *minutos de jogo*.

Após chegarem a um consenso acerca das coordenadas destes pontos, a professora partiu para a construção, no quadro, da expressão analítica da reta que passava pelo ponto indicado pelo Leonardo e pelo centro de gravidade: “nós, no primeiro período, aprendemos a construir retas!” (observação 6). Pedindo a colaboração dos alunos, chegaram à equação: $y=0.36x+1.58$.

De seguida, a professora deu instruções para que esta reta fosse desenhada na calculadora gráfica através do menu *DrawF (equação da reta):*

> **Professora:** O que é que aconteceu? Para quem já desenhou essa e [já] tinha uma outra desenhada, são muito diferentes?
> **Vários alunos:** Não.
> **Jaime:** Não, [no meu caso] elas intersetam-se.
> **Professora:** Então parece que se cruzam. Pronto… (…) Ora bem, desenharam a reta que pedi… mas sabem uma coisa? A máquina [calculadora gráfica] faz isso tudo sozinha! Viram esta trabalheira que estivemos a fazer (…) Agora quero que vocês comparem… a que obtiveram pela máquina [reta de regressão] com essa à mão [$y=0.36x+1.58$] em que se determinou a expressão analítica?... Sabem como é que se faz na máquina? Para quem já fez na máquina, digam-me lá como é que se faz?
> **Alunos:** [há alunos que dizem que sim, outros dizem que não sabem fazer na calculadora gráfica]
> **Isabel:** Mas, ó Stora, mas não dá valores iguais! [referindo-se ao facto dos declives e ordenadas na origem não serem iguais nas duas retas]
> **Professora:** Pois não! … É assim, eu disse-vos um ponto que tinha a certeza que a reta da máquina [regressão] passava nele [centro de gravidade], os vossos colegas disseram o outro [(22,4; 9,6)] … e se ele não passa por (22,4; 9,6)?
> **Professora:** [dá instruções para se chegar à reta de regressão através da calculadora gráfica, explicando cuidadosamente o que cada uma delas faz] E agora a equação [regressão linear] que a máquina me deu foi $y=0.43x+0.135$. Concordam?
> **Ana:** Concordo.
> **Professora:** E agora façam *graph* [para aparecer no visor a reta de regressão linear]... Ficou muito diferente da nossa [reta estimada]?
> **Vários alunos:** Nem por isso. Não. (observação 6)

Na exploração desta tarefa, a professora procurou inicialmente que os alunos sugerissem ou indicassem a equação de uma reta que é a que *melhor se ajusta aos pontos do diagrama de dispersão* de coordenadas (minutos de jogo; pontos obtidos). Isto porque tinha a intenção de lhes pedir para compararem esta reta sugerida ou determinada por eles com a que poderia ser obtida de modo expedito na calculadora gráfica, a reta de regressão, através do uso da instrução *LinReg(a+bx).* Na aula, quando confrontaram essas duas retas, chegou-se à conclusão que elas eram diferentes mas que não estavam muito afastadas entre si quando visualizadas em simultâneo sobre o diagrama de dispersão na calculadora gráfica.

A professora reagiu ao facto de a aluna Isabel ter ficado surpreendida por essas duas retas não apresentarem o mesmo declive e ordenada na origem explicando que, dos dois pontos que se tinham usado para determinar a equação da reta, só conseguia garantir que um deles, o centro de gravidade, era mesmo um ponto da reta de regressão procurada. Nessa interação acabou por ser assumido que a reta de regressão é aquela que melhor se ajusta à nuvem de pontos considerada e que o centro de gravidade é um ponto que lhe pertence, sendo que este último facto foi confirmado quando os alunos verificaram que as coordenadas do centro de gravidade satisfaziam a equação da reta de regressão (da calculadora gráfica). No entanto, no desenvolvimento desta tarefa, não surgiram outras explicações que ajudassem a perceber com mais profundidade esses factos assumidos.

Tal como aconteceu com a aluna Isabel, durante o desenvolvimento da tarefa, houve ainda alguns alunos que se mostraram intrigados com o facto de as duas retas que determinaram não serem a mesma. A professora procurou clarificar novamente esta situação:

> (…) Vocês viram um ponto que [achavam que] estava na reta… Como vocês não viram mais nenhum, eu sugeri um outro [ponto]… Calculamos à moda antiga a reta[ou seja, obteve-se à mão a expressão da reta que passava pelos dois pontos indicados], desenhei-a à mão… Depois fomos ver se o meu desenho [representação desta reta inicial, na calculadora gráfica] estava muito afastado ou não do desenho [reta] que a máquina fazia [reta de regressão linear] e vimos que quanto ao declive até nem estava muito mal [com uma diferença de cerca de 7 centésimas, considerando estes valores próximos], já a ordenada na origem calhou um bocado

> mal… Mas isto à mão!... Com as opções tomadas que foram resultado da observação do gráfico de dispersão… (observação 6)

Nesta explicação, Estela acabou por descrever novamente a sequência de trabalho realizada na aula de modo a obter-se a reta estimada. Realçou igualmente o facto de a equação desta reta não ser coincidente com a da regressão, que se obteve através da calculadora gráfica. Explicou ainda que essa situação seria altamente improvável uma vez que a estratégia usada por ela e pelos alunos na construção da *reta que melhor se ajusta aos pontos* houve seleção de pontos a partir da observação do gráfico e este não foi o método usado pela calculadora. Estela optou por não mencionar o método dos mínimos quadrados, mas explicou aos alunos que a reta de regressão só se poderia obter na calculadora gráfica pois era precisamente a reta que procuravam com a melhor precisão possível.

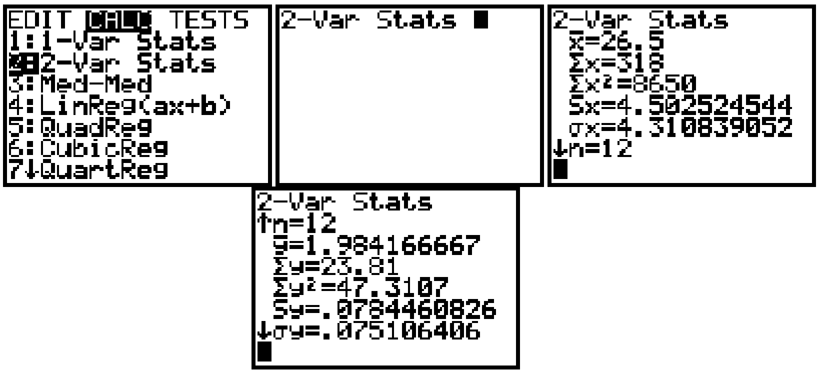
Centro de gravidade

O centro de gravidade foi mencionado pela primeira vez na tarefa introdutória do tema distribuições bidimensionais (t16), assim que os alunos se depararam com um conjunto de instruções que a professora forneceu de modo a poderem usar a calculadora gráfica para representar os dados bivariados no diagrama de dispersão e realizar o cálculo de medidas estatísticas da distribuição bidimensional em estudo (figura 37).

C. Cálculo da média de duas variáveis em simultâneo na calculadora

(STAT + CALC + opção2: 2-Var Stats):

Calculo em simultâneo, por exemplo, da média das ***idades*** e da média das ***alturas*** dos jogadores do Porto. Como os dados dos quais pretendemos calcular a média estão, respetivamente, na lista 1 e na lista 2, basta só dar uma instrução à máquina.



Vendo esta sequência de écrans na calculadora, podemos dizer que a média das idades é 26.5 anos e que a média das alturas é 1.98 metros.

Ao ponto $(\overline{x}, \overline{y})$ chamamos **centro de gravidade** da distribuição bidimensional.

Figura 37: Instruções para o cálculo das médias.

O centro de gravidade foi também calculado pelos alunos nesta tarefa introdutória (t16), quando foi usado para se chegar à expressão analítica da reta que parecia melhor descrever a relação bivariada expressa no diagrama de dispersão. Foi também calculado com os dados da tarefa *Massa corporal feminina* (t21). O centro de gravidade foi tomado como o ponto que tem como abcissa a média dos dados referentes a uma variável, e como ordenada a média dos dados da outra variável, determinados na calculadora gráfica através do menu *STAT* (*CALC* e opção *2-Var Stats*). Estela também assumiu o centro de gravidade como um ponto que pertence sempre à reta de regressão. Ao contrário do que fez com a tarefa 16, no trabalho em torno da tarefa 21, a professora explicou e mostrou, através da calculadora gráfica, como proceder à marcação do centro de gravidade sobre a reta de regressão quando esta está exibida no ecrã da calculadora, com *Draw* (opção *draw point*) e *CALC* (opção *value*).

Nas aulas de Estela, este conceito foi sobretudo tratado numa perspetiva processual e, embora tenha sido relacionado com a reta de regressão, não foi desenvolvido ao nível concetual.

Coeficiente de correlação linear

Tarefa *A equipa de basquetebol do Porto* (t16)

Para introduzir a ideia de correlação entre as variáveis *tempo* e *pontos*, houve a seguinte interação:

> **Professora:** Então, prestem atenção aqui a uma outra coisa (observando o diagrama de dispersão)… À medida que o tempo aumenta… que os jogadores estão mais tempo em campo, em geral, eles marcam mais pontos… certo? Então, por essa razão, dizemos que há correlação linear positiva… certo?… E quanto mais esses pontos se aproximam da reta… quanto menor for a distância dos pontinhos [do diagrama de dispersão] à reta de regressão… for menor para todos eles, mais forte é essa correlação!
> **Dalila:** É o que tu estavas a dizer (para o Leonardo).
> **Leonardo:** São aqueles quadrados…
> **Professora:** São aqueles quadradinhos a ficar mais pequenos (referindo-se aos resíduos/desvios que o aluno Leonardo tinha descoberto na sua *TI-Nspire*)… Como é que se mede essa correlação matematicamente?… se é forte ou se é menos forte? À custa de cálculos que a máquina faz….Claro que vos podia dar a fórmula, aliás vocês têm a fórmula no livro… (observação 5)

Estela mostrou, adicionalmente, a fórmula do coeficiente de correlação da página 81 do manual e, ainda, a síntese geral dos valores que o coeficiente de correlação pode tomar da página 82. Leu também a seguinte definição do coeficiente de correlação, incluída no enunciado da tarefa, e que estava em consonância com a síntese do manual: "O coeficiente de correlação é um número que mede a correlação entre duas variáveis. Representa-se por r e varia entre -1 e 1. Se a correlação é positiva, então $r>0$. Se a correlação é negativa, então $r<0$. Quanto maior o valor absoluto de r, maior é a correlação. Quando $r=0$, não existe correlação" (observação 5).

No excerto acima, a professora tentou associar algumas ideias: (1) a tendência global observada no gráfico de dispersão para um aumento do *número de pontos* à medida que os *minutos de jogo* aumentam, com o declive positivo da reta de regressão e ainda com *correlação linear positiva*; e (2) a maior proximidade da reta de regressão com os pontos do diagrama de dispersão traduzida como uma forte correlação linear existente entre as variáveis *minutos de jogo* e *pontos obtidos*.

Apesar de a professora mencionar, em voz alta, que os "quadrados" (que o Leonardo tinha conseguido fazer na calculadora gráfica) teriam de ficar cada vez mais pequenos para a correlação linear ser cada vez mais forte, a maioria dos alunos não se apercebeu efetivamente do que tinha sido feito por este aluno, dado que não lhes foi mostrado o ecrã. No entanto, este seu comentário revela algum entendimento acerca da relevância da análise residual na avaliação da força da associação linear entre duas variáveis consideradas.

Através desta tarefa, Estela indicou adicionalmente como se ativar o *r* (símbolo associado ao coeficiente de correlação) na calculadora gráfica, de modo a que este valor aparecesse automaticamente sempre que se determinasse a reta de regressão. Assinalando, também, como o determinar na calculadora gráfica, foi feita a seguinte leitura do resultado de *r* obtido:

> **(**Vários alunos dizem ter obtido um coeficiente de correlação de aproximadamente 0.9).
> **Professora:** Quanto mais próximo de 1, mais forte é a correlação. Também o é quanto mais próximo de -1. Só que, neste caso (*r*=0.9), é uma correlação forte positiva… Porquê?... Porque a reta [de regressão] tem uma inclinação com declive positivo. Se o *r* for negativo, tem uma reta de regressão com declive negativo… e quanto mais próximo de -1, mais forte. Está?
> **Ivo:** Stora, não consigo ver o *r* na minha calculadora (não o sabe instalar).
> **Professora:** Espera que já vou aí ver… a tua máquina deve ser especial…(ouvem-se risos). (observação 5)

O valor de *r* de 0.9 foi descrito pela professora quanto à força, ou seja, quanto mais forte, mais próximo de -1 ou de 1, e quanto ao sinal, negativo ou positivo. Este sinal foi também associado ao declive da reta de regressão. Neste excerto, também é possível observar que há um aluno que acusou não estar a visualizar o valor de *r* na sua calculadora depois de ter efetuado o procedimento para obter a equação da reta de regressão.

Tarefa *Associação entre nuvens de pontos e coeficientes de correlação* (t17)

A professora escolheu, do manual, algumas tarefas com um conjunto de gráficos de pontos ou diagramas de dispersão aos quais deveria ser associado os respetivos valores de coeficiente de correlação linear de um conjunto de valores disponíveis. A primeira tarefa trabalhada na aula foi a tarefa 17, que foi apresentada do seguinte modo: "Só por observação, sem saber o que estas variáveis representam [na primeira tarefa deste tipo, os dados estão representados num quadriculado que não tem os eixos nomeados], devem relacionar cada nuvem de pontos, com os coeficientes de correlação" (observação 6).

Ao oferecer algumas sugestões para abordar esta tarefa, Estela indicou aos alunos que, nos seus julgamentos, poderiam considerar a reta que *melhor se ajusta aos pontos* de cada diagrama de dispersão, e aquela que melhor satisfizesse esta condição era a que corresponderia a um coeficiente de correlação mais forte. Estela também sugeriu uma outra estratégia inicial: atribuir uma escala à quadrícula, determinar as coordenadas de cada ponto do diagrama de dispersão, colocar esses dados na calculadora gráfica e, através dela, obter a expressão analítica da reta de regressão, bem como o valor do coeficiente de correlação. Como a maioria dos alunos indicou que não estava a perceber como proceder seguindo as indicações da professora, Estela resolveu demonstrar o que fazer. Começou por pedir à investigadora para lhe ditar os valores das abcissas (lista 1) e das ordenadas (lista 2) do gráfico 4 (ver anexo 3) (tendo, para o efeito, indicado a escala desejada). Enquanto se preparava para inserir os valores no computador, no *software* da calculadora gráfica que estava a ser projetada sobre o quadro, disse:

> **Professora:** Sandra, que valores tens para a lista 1?
> **Investigadora:** 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.
> **Professora:** E para a lista 2?
> **Investigadora:** 9, 8, 8, 6, 6, 5, 6, 4, 5, 3.
> **Professora:** [Mostra no ecrã o diagrama de dispersão com a reta de regressão, y= -0.58x+9.2, incluída e determina também o valor do coeficiente de correlação para estes dados, o qual é de aproximadamente -0.93]. (observação 6)

Com esta estratégia inicial, obteve-se um valor coeficiente de correlação de 0.93 que não coincidia exatamente com o correspondente valor do enunciado que era de -0.94; contudo, esta situação não foi mencionada. A professora também mostrou como é que os alunos poderiam chegar à equação da reta de regressão e ao coeficiente de correlação, acabando por não apelar à interpretação da correlação.

Os alunos, de uma maneira geral, não revelaram dificuldades em realizar as associações corretas entre valores do coeficiente de correlação respetivos diagramas, que foram sempre validadas pela professora. Contudo, ao longo dessas interações, a professora nem sempre lhes pediu para justificarem as suas respostas.

Tarefa *Associação entre diagramas de dispersão e coeficientes de correlação* (t18)

É uma tarefa semelhante à anterior. Depois dos alunos indicarem corretamente as associações, a professora resolveu interrogá-los sobre a questão do sinal da correlação por se ter apercebido da existência de alguns alunos com dificuldades nessa parte, havendo a seguinte interação:

> **Professora:** Ora bem, eu, pelos gráficos (ver figura?), vejo que haverá dois que apresentam uma correlação positiva e outros dois negativa. Porque é que eu estou a dizer isso? … É que ouvi aqui… assim um burburinho… a comentar assim…Quais os que apresentam correlação negativa? O I e o IV, certo? (…) (dirigindo-se à aluna Márcia, diz) Achas que se, traçando uma reta que se aproxime… que pode não passar por todos os pontos… que se aproxima o máximo possível por todos os pontos… Achas que essa aproximação é maior no gráfico I ou no IV?
> **Márcia:** É no I.
> **Professora:** Muito bem, então quais dos dois valores negativos, o -0.3 e -0.8 é que é o mais forte? (ouve-se uma outra aluna, Sara, a responder).
> **Sara:** Acho que é o [valor] mais perto de 1.
> **Márcia:** É o -0.3
> **Professora:** (Não dizendo que discorda com a resposta dada, continua) Então… Pensa numa reta graduada. Tens 0, o 1 e o -1. Se os coeficientes são negativos… Qual deles é que está mais perto de -1…? (Desenha uma reta no quadro com os pontos -1, -0.8, -0.3, 0 e 1 marcados).
> **Márcia:** Ah, é o -0.8!
> **Professora:** Está [entendido]?! (observação 6)

Neste excerto, observa-se então que há alunos a fazer a associação correta entre correlação negativa "mais forte" e o diagrama de dispersão I (ver figura 38); contudo, o mesmo não estava a acontecer entre esta associação e a quantificação numérica da correlação. Estela tentou ajudar essas alunas a entender que, no caso de correlação negativa, o coeficiente de correlação mais forte seria o de valor mais próximo de -1 e não o mais próximo de 1, como elas estavam a indicar erroneamente.

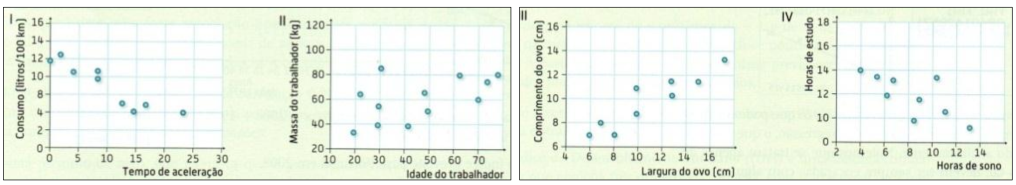


Figura 38: Gráficos I, II, II e IV (t18).

Adicionalmente, com os mesmos gráficos e valores de coeficiente, foi pedido aos alunos para descreverem a correlação entre as duas variáveis observadas em cada diagrama, tal como a segunda parte do enunciado requeria. A descrição realizada destaca-a pela "força" (forte ou fraca) e pelo "sinal" da relação (positiva ou negativa):

> **Professora:** (lê no enunciado da tarefa) 'Descreva a correlação entre as variáveis de cada diagrama'. Pedro, no diagrama I, [observando] aquela nuvem de pontos, existe correlação…?
> **Pedro:** Positiva…
> **Professora:** Positiva…?!
> **Pedro:** Estava a brincar! (vários risos) É negativa, professora…
> **Professora:** Forte ou fraca?
> **Pedro:** Forte porque está perto de -1.
> **Professora:** Ah! (indica que concorda, abanando a cabeça). Agora para o diagrama II… Diz lá Sara…
> **Sara:** A correlação é positiva e fraca, porque o valor é afastado de 1.
> **Professora:** Quem diz para o III?
> **Dalila:** Correlação é positiva e forte.
> **Professora:** Para o IV. Quem diz?
> **Lurdes:** A correlação é negativa e… acho que é fraca.
> **Professora:** Porque?… Qual o valor de coeficiente de correlação que lhe associaste?
> **Lurdes:** é o -0.3.
> **Professora:** Então é fraca! [Já agora] Como é que se chama à correlação zero?
> **Alguns alunos:** Nula! Chama-se nula… (observação 6)

Nesta interação, observa-se que foi usado o argumento que, para valores de coeficientes de correlação linear próximos de 1 ou -1, a correlação seria forte, e que, para valores mais próximos de zero, seria fraca. Este argumento está em sintonia com a definição de coeficiente apresentada na aula.

Tarefa *Estudo com carateres qualitativos e quantitativos* (t20)

Depois de recolhidos os dados na turma, a professora deu indicações para que os alunos partissem para a "organização dos dados… trabalhassem esses dados" (observação 7) com as representações que quisessem. A uma dada altura, pediu para que os alunos se concentrassem mais no estudo da relação bivariada *alturas* e *pesos* (dos alunos da turma): "Como não há muito tempo… para agora quero que verifiquem a existência de correlação… Queria que fizessem o diagrama de dispersão e verificassem se há ou não correlação entre altura e peso na vossa turma" (observação 7).

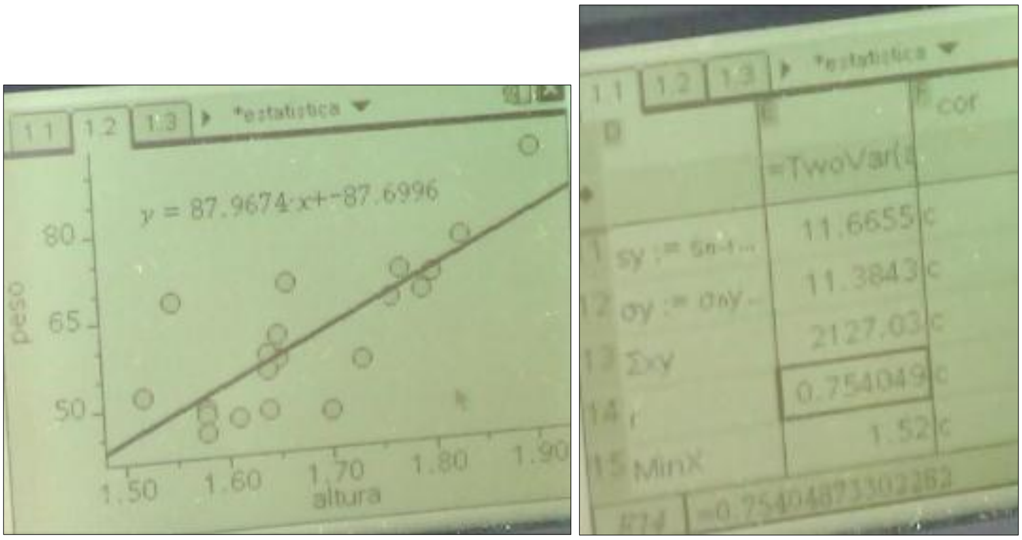


Figura 39: Ecrãs da calculadora gráfica (diagrama de dispersão; folha de cálculo que assinala um valor de *r* de aproximadamente 0.75).

A pedido da professora, a aluna Dalila mostrou os dois ecrãs da calculadora gráfica à investigadora (figura 39). Lurdes, outra aluna, por sua vez, foi ao quadro assentar alguns resultados obtidos de estatísticas sumárias (ver figura 40) para todas as variáveis consideradas no enunciado. Incluiu, em particular, a média das duas variáveis (*altura* e *peso*) em que se estava a averiguar a associação linear. Era suposto que esta aluna tivesse apresentado no quadro mais resultados, nomeadamente, relativos à relação bivariada solicitada pela professora no decurso da aula. Contudo, a aula terminou sem que ela tivesse conseguido registar tudo o que pretendia no quadro.

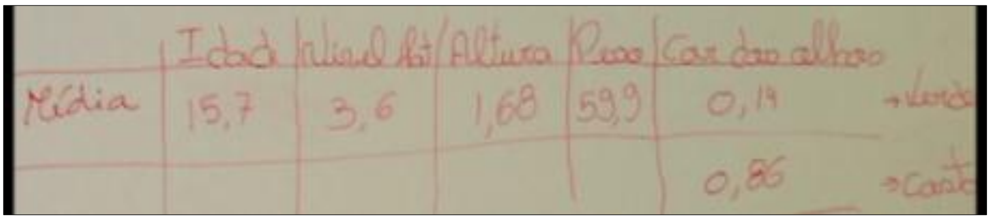


Figura 40: Excerto do registo no quadro de Lurdes (resultados de média para diferentes categorias).

5.3.5.2 O modelo de regressão linear

De seguida, descreve-se e analisa-se atenção dada ao modelo de regressão linear através das tarefas *População residente em Portugal* (t19) e *Massa corporal feminina* (t21).

Tarefa *População residente em Portugal* (t19)

Quando propôs esta tarefa, a professora foi solicitando aos alunos que introduzissem as listas de dados na calculadora gráfica e, através dela, obtivessem o diagrama de dispersão e a reta de regressão linear (tal como a que estava exibida no enunciado (t19)). Pediu ainda que determinassem o valor do coeficiente de correlação linear das duas variáveis em causa.

Ao pedir para analisar a correlação graficamente, a professora e alguns alunos pronunciaram-se quanto à existência de correlação e classificaram-na quanto à força (forte ou fraca) e ao sinal (correspondente ao sinal do declive da reta de regressão):

> **Professora:** O que vos parece? Existe correlação ou não existe?
> **Professora e vários alunos:** Existe.
> **Professora:** Como é que a descrevem?... É fraca?
> **Vários alunos:** É forte.
> **Professora:** Ah… Mas é positiva ou negativa?
> **Vários alunos:** Positiva
> **Professora:** É positiva, porquê?... Já não me lembro!…
> **Ivo:** Porque o declive da reta é positivo. (observação 7)

A existência de correlação foi descrita pela força através da palavra “forte”, mas não foi explicado o critério usado. Pensa-se que terá sido por Estela e os alunos considerarem que a reta de regressão (que, na situação concreta, não era necessário ser imaginada, como habitualmente, pois vinha incluída no diagrama de dispersão do enunciado) estaria próxima da maior parte dos dados bivariados.

Depois de dar algum tempo para os alunos trabalharem nesta tarefa, Estela pediu a um aluno que fosse ao seu computador para trabalhar na calculadora (que já estava exibida na tela sobre o quadro branco) todas as representações que tinham sido pedidas. Como o aluno (Mário) manifestou que não tinha muita prática com a calculadora em exibição, a professora pediu a uma outra aluna (Diana) para o apoiar. Estes dois alunos chegaram à mesma equação da reta de regressão indicada no enunciado e ainda determinaram aproximadamente os valores do coeficiente de correlação (r=0.988) através da instrução *LinReg* da calculadora gráfica. A figura 41 mostra essa projeção:

Figura 41: Expressão analítica da reta de regressão e valor de r.

Nesta tarefa, depois do valor de coeficiente de correlação ter sido determinado na calculadora gráfica, foi também interpretado por alguns alunos. Estela escreveu no quadro a resposta que pretendia que ficasse registada nos cadernos (figura 42), dizendo em voz alta: “A correlação é positiva e forte pois… Exatamente, eu vou escrever o que disseste Ivo!… o seu valor está próximo de um. E com três casas decimais é 0.988.” (observação 8), acabando por fazer uma leitura semelhante à correlação observada no gráfico do enunciado.

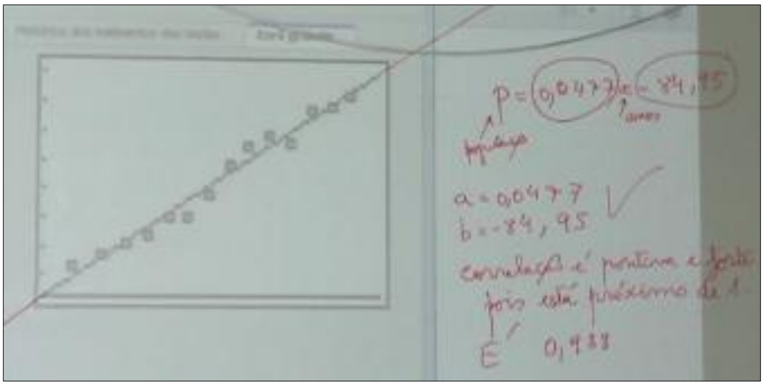


Figura 42: Projeção do diagrama de dispersão e registo à mão da expressão analítica da reta de regressão e do valor de *r*.

Através desta tarefa, a professora refere-se à reta de regressão como um modelo linear que pode ser usado para fazer estimativas ou previsões. Para ajudar os alunos a pensar sobre a inadequação da reta de regressão para se estimar ou fazer previsões a médio e longo prazo (tal como a segunda alínea requeria), Estela teve a seguinte interação com os alunos:

> **Professora:** Portanto… Qual o papel principal deste modelo linear, ou seja, desta reta de regressão? O que ela serve é para estimar, fazer uma estimativa… prever! E o que se está aí a afirmar é que este modelo não serve para estimar… isto aqui são os anos que metemos na máquina, certo? Aliás, se fizerem o teste, dá aqui cada um dos anos e a respetiva população que havia, certo? O que se está aí a afirmar é… Porque é que este modelo não serve para eu imaginar qual será a população daqui a não sei quantos séculos, nem serve para imaginar quantas pessoas existiam há não sei quantos séculos atrás?… Ah!... Tenho ali já um candidato (apontando para o aluno que está com um braço levantado).
> **Ivo:** Porque o desvio-padrão não é 1 e os anos não são diretamente proporcionais ao número de habitantes.
> **Professora:** Esta é uma explicação puramente matemática, é de Matemática A. É verdade… É a diferença [de resposta] entre os alunos da Matemática A e os alunos de MACS… Outro aluno? (observação 7)

Nesta interação, a professora tenta fazer com que os alunos se apercebam da limitação do modelo linear na previsão a médio e longo prazo da evolução da população portuguesa ou na estimativa de um valor aproximado da população portuguesa nesses períodos. Quanto à resposta fornecida pelo aluno em que este tenta dar como explicação sobre a restrição do modelo linear, a professora não o questionou para entender porque é que ele mencionou o desvio-padrão, talvez por considerar que tinha sido um lapso e que ele pretendia referir-se ao coeficiente de correlação linear. Também não o questionou por ter apontado a não existência de proporcionalidade entre as variáveis *anos* e *população portuguesa* (em milhões).

A discussão em grande grupo continuou nos seguintes moldes:

> **Professora:** Quem diz mais alguma coisa?
> **Mariana:** Ao substituir-se o *a* por um ano [na equação da reta de regressão]… e a população dar um valor normal.
> **Professora:** O que é a população dar um valor normal?
> **Leonardo:** Superior a zero
> **Professora:** Superior a zero, pelo menos… Diz mais alto [disse para o Leonardo por ele ter falado muito baixo]?
> **Leonardo:** Ao prolongar-se a reta [esquerda] vai passar por baixo de zero.
> **Professora:** Exatamente. Por exemplo, se prolongarmos a reta [à esquerda] o que acontece? [ver na figura 43 o mesmo prolongamento acrescentado pela professora a marcador vermelho ao gráfico inicial]
> **Alunos:** Tínhamos população negativa.
> **Professora:** Isto é impossível! Portanto há uns séculos atrás teríamos população negativa… Em contrapartida, se prolongássemos a reta [à direita]? O q acontecia?
> **Alunos:** A população vai crescer.
> **Professora:** A população aumentava. A população crescia infinitamente. Isto não é possível? Porquê? Está aí [no manual] uma sugestão.
> **Jaime:** Não cabíamos cá todos…
> **Rafael:** O mundo vai acabar…

Professora: Não cabíamos cá todos. Mais? Os recursos são…?
Professora e alunos: Limitados
Ivo: Há aqui qualquer coisa que está mal. Era para dar assim uma resposta dessas ou…?
Professora: [a aluna Maria pergunta à professora qual tinha sido a resposta do Ivo] Ivo repete, por favor, o que tinhas dito há bocado.
Ivo: O quê?... Eu disse…
Professora: Ah! Exatamente… Disse que o crescimento da população não é diretamente proporcional ao… passar do tempo. Como ele estava a ver que o coeficiente de correlação até é muito próximo de 1 e estava a ver que havia uma certa proporcionalidade e até lhe parecia que passava aqui pela origem. Portanto ele até previu uma certa proporcionalidade direta, não foi? Pronto… Previu, mas isso são tudo conjeturas. Está?! Está percebido? Pronto. (observação 7)

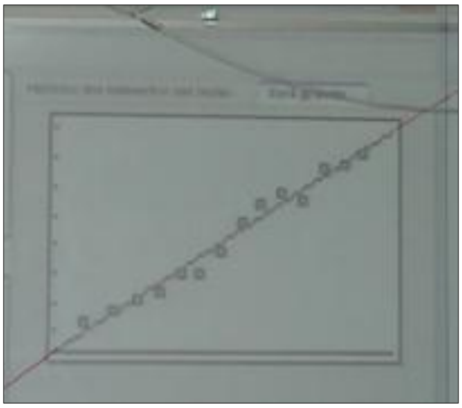


Figura 43: Diagrama de dispersão.

Pensa-se que, pelo facto de Estela ter destacado esta resposta do aluno Ivo durante a interação, fez com que essa mesma resposta surgisse novamente na discussão. Desta segunda vez, é a professora que acaba por reproduzi-la e depois justificá-la. E quanto à questão da não proporcionalidade direta observada pelo Ivo na reta de regressão, a professora apresentou argumentos para defender uma possível proporcionalidade direta entre as variáveis em estudo que o gráfico de dispersão (ver figura 43) lhe parecia mostrar visualmente (dado a reta de regressão no seu prolongamento à esquerda parecer passar na origem). Apesar de ter sido falado na sala de aula que a reta não poderia ser usada, por exemplo, para estimar a população de há uns séculos, por motivo desta estimativa dar uma população negativa, o uso da reta de regressão na concretização de valores específicos poderia ter facilitado o entendimento da limitação do modelo linear na previsão ou estimativa a médio e longo prazo, bem como uma melhor perceção dos dados, da sua evolução e do seu contexto. Nesta tarefa, a professora fez transparecer a ideia que o modelo linear era totalmente desapropriado em vários momentos no tempo, mas não deu exemplos concretos e não evidenciououtras situações em que a sua aplicabilidade poderia ser particularmente útil.

Tarefa *Massa corporal feminina* (t21)

Esta tarefa, selecionada de um manual, envolvia 15 dados bivariados que correspondiam à altura (em centímetros, com valores a variar entre 157 e 172 cm) e massa corporal (em kilos, com valores a variar entre 52 e 72 kg) de mulheres. Os alunos teriam de usar a reta de regressão para estimar a massa corporal de uma mulher com 175cm altura.

Esta proposta foi praticamente toda trabalhada na calculadora gráfica que se encontrava projetada na tela. Enquanto os alunos tentavam usar as suas calculadoras para resolver a tarefa, a professora trabalhava na calculadora projetada. Assim, foi possível observar Estela a colocar as listas na calculadora. Através da instrução *LinReg* chegou rapidamente à equação da reta de regressão e ao valor do coeficiente de correlação linear (que também registou no quadro, para ficar com o diagrama de dispersão e a informação determinada à vista em simultâneo – ver figura 44):

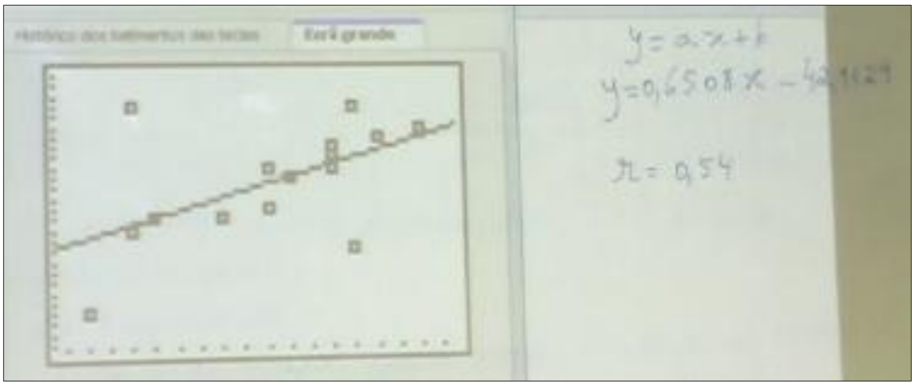


Figura 44: Diagrama de dispersão com a reta de regressão e registo no quadro da equação da reta e do valor de *r*.

Para se estimar o peso de uma mulher com 175 cm de altura, houve um aluno (Jaime) que sugeriu que se usasse a instrução *value* do Menu *Calculate* da calculadora e se colocasse como abcissa *x=175*, para a calculadora dar o respetivo valor de *y*. A calculadora deu erro com essa instrução e professora fez notar que o problema deveria estar na janela de visualização, cujo *Zoom* (*ZoomStat)* deveria estar em concordância com a amplitude dos dados e que, por isso, não poderia incluir valores superiores a 172 cm: "É a janela… *Window*!... Nas listas [dos dados na calculadora gráfica] não há valores superiores a 172, pois não?" (observação 8). Estela não se lembrou que bastava mudar de *Zoom* (por exemplo, para *ZStandard*) e ajustar as suas dimensões para se determinar a massa corporal desejada com a sugestão do Jaime. Como alternativa, a professora sugeriu que os alunos usassem o menu tabela (*Tblset*, ver figura 45) da calculadora para determinar a ordenada procurada: "Não dá para usar a sugestão do Jaime… Já só dá com a tabela [*2nd Tblset* e indicar em *Tblstart* o valor 175]. Perceberam bem?" (observação 8).

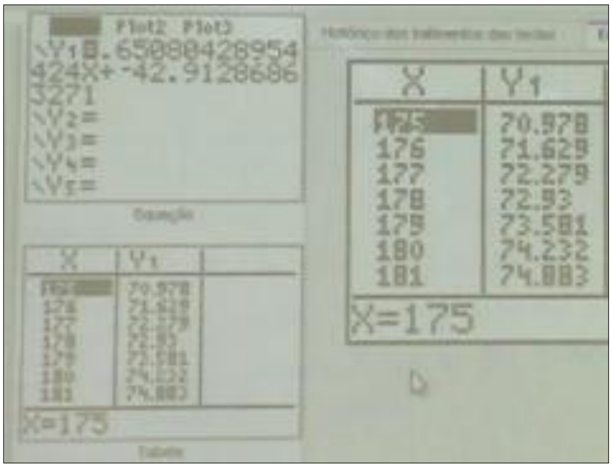


Figura 45: Tabela que mostra a concretização de alguns valores de *x* (da reta de regressão *y* em função de *x*).

Ao frisar a utilidade do modelo linear nesta situação, Estela parece atribuir-lhe pouca credibilidade: "Estou a usar a equação da reta. E o modelo não é que seja verdade! Serve para fazer uma estimativa… A partir dele vou ver mais ou menos… qual será [o peso] de uma pessoa com 175 cm de altura, está bem?" (observação 8). Neste comentário apesar de a professora ter notado que o valor da variável independente fornecido, 175, não pertencia ao intervalo de variação dos restantes valores da variável *altura*, não fez qualquer comentário ao facto de, habitualmente, só se poder estimar com valores da variável independente dentro e numa vizinhança muita próxima deste intervalo. A pouca fiabilidade atribuída ao modelo poderá estar associada ao facto do valor do coeficiente de correlação, desta vez, rondar os 0.5, um valor moderado para a força da correlação; contudo, não surgiu nenhum comentário nesse sentido.

5.3.5.3 Reflexões sobre o tópico na aula

Estela considerou pertinente, neste tema, como atividade inicial, colocar os alunos à procura de uma reta que melhor descrevesse a relação bivariada exposta no diagrama de dispersão com dados da tarefa *A equipa de basquetebol do Porto* (t16), sem que os alunos soubessem ainda como determinar, na calculadora gráfica, a reta de regressão. Ela pretendia que os alunos chegassem às retas (e suas expressões analíticas) com as suas estratégias e posteriormente levá-los a "comparar uma das equações de reta obtidas [por eles] com a equação da reta de regressão" (entrevista 3), que lhes iria ensinar através da calculadora.

Esta atividade, por um lado, envolveria também que os alunos utilizassem os seus conhecimentos prévios sobre funções, adquiridos no 2.º período (em particular, para obterem a expressão analítica da reta a partir da indicação de dois pontos observados) e, por outro, poderia facilitar que eles desenvolvessem as suas ideias iniciais sobre regressão linear, nomeadamente, ao analisar o quão afastada ou próxima a reta estimada se encontrava da reta de regressão (sobretudo graficamente).

Ainda quanto ao tópico da regressão linear, na entrevista, Estela lamentou não ter aproveitado os "quadrados" que o aluno Leonardo descobriu na sua calculadora (*TI-Nspire*) para desenvolver mais este tópico junto dos alunos: "Foi uma coisa que nunca falei na regressão linear… mas se calhar… com mais tempo podia ter explorado esses quadrados [na calculadora gráfica] que tinham a ver com desvios… Se calhar com esses alunos teria sido interessante" (entrevista 3).

A breve referência que fez na aula sobre "os quadrados" não foi captada pela maioria dos alunos visto que este trabalho do aluno Leonardo não foi mostrado à turma. Estela revelou entender a ideia principal implícita na construção da reta de regressão quando disse, a seu respeito, que "a soma dos quadrados dos desvios é mínima" (entrevista 3).

Quando questionada, na entrevista, sobre como é que o centro de gravidade tinha surgido na aula, Estela confessou que a ideia que este ponto pertencia à reta tinha sido indicada por algum aluno que provavelmente verificou as coordenadas através da equação da reta de regressão. Mas, a partir do momento que fez a verificação deste facto com os alunos na calculadora, Estela não sentiu necessidade de o desenvolver mais. E não tem dúvidas que:

> Os alunos têm noção que este ponto [centro de gravidade] pertence à reta [de regressão] (…) Ensinei a ver na calculadora… Se escrever *x*= com o valor da média [com as instruções: *2nd*, *calc*, *opção 1: value* sobre o ecrã com reta de regressão e diagrama de dispersão exibidos], dá o valor da média do *y* e simultaneamente dá para ver que estava lá o ponto [visualmente, sobre a reta de regressão]. (entrevista 3)

Quanto à aplicabilidade da reta de regressão, Estela traduziu-a como “um modelo a curto prazo” e, ao comentar sobre a perceção que os alunos ficaram sobre esse modelo, refere:

> (…) uma coisa são as nuvens de pontos, os dados que observaram [no diagrama de dispersão] e outra coisa é um modelo, uma reta de regressão… Esta reta serve para fazer estimativas, mas são estimativas… não são realidades!... [que muitas vezes] podem não corresponder ao real… (…) eles [alunos] ficaram a saber isso. (entrevista 3)

Ao explicitar o que entende pela “não realidade” dos valores das estimativas efetuadas através de retas de regressão, Estela apontou precisamente a tarefa *População residente em Portugal* (t19) proposta na aula (que serviu para mostrar algumas restrições no uso do modelo linear para estimar a médio e longo prazo) recordando que:

> … os alunos até disseram que com o modelo linear [desta tarefa] a população dava negativa [há uns quantos séculos atrás] e isso não é real e a muito longo prazo disseram que era impossível [usá-lo] porque, de acordo com a realidade, os recursos naturais vão diminuindo e a população não cresce indefinidamente como o modelo sugere. (entrevista 3)

A professora deu a entender que o modelo linear era útil para estimar num determinado intervalo, embora não especificando que esse intervalo deveria coincidir com o intervalo de variação dos dados da variável independente. Ainda sobre o modelo linear, observa-se que clarificou que ele não deveria ser utilizado para se fazer previsões a longo prazo.

Na exploração da tarefa *Associação entre nuvens de pontos e coeficientes de correlação* (t17), Estela sugeriu o recurso à calculadora gráfica como uma das primeiras abordagens. Na entrevista, quando questionada sobre esta opção, Estela relembrou que os alunos ficaram sem saber o que fazer e que, assim, criou mais uma oportunidade para lhes mostrar a utilidade da calculadora para explorar os dados e para obter resultados:

> Para eles próprios verem que também podem introduzir na calculadora gráfica os dados… [ou até dados de outras] situações do manual… Se tiverem dúvidas e não conseguirem ver mais ou menos de cabeça, fazer uma estimativa [aventar uma possível resposta, por observação] (…) Podem experimentar [com a calculadora, não precisam de ficar parados] sem problemas de discutir, errar, … Aqui não há errar… há experimentar e concluir, colocar hipóteses e confirmá-las ou não. (entrevista 3)

Na perspetiva da professora, os alunos deveriam explorar e experienciar diferentes estratégias na resolução de tarefas, destacando que na exploração de uma tarefa a calculadora gráfica pode ter um papel importante como abordagem inicial.

A professora reconheceu a possibilidade de se propor tarefas para a aula que peçam para se encontrar outros modelos que não o linear pois teria sido interessante para os alunos verem a proximidade da Matemática e da Estatística, mas "o tempo não é muito e o programa não pede isso… Não se perde tempo com dados que possam não ter uma relação que não seja linear… Os manuais não fazem isso" (entrevista 3). Para Estela, de uma maneira geral, a averiguação da existência de associação linear entre duas variáveis passa pela observação do diagrama de dispersão, pela determinação da reta de regressão e pelo cálculo do valor do coeficiente de correlação linear. Quando esse último valor é fraco assume que a relação entre as variáveis não é certamente linear, mas pode ser "quadrática, cúbica, etc…" (entrevista 3).

5.3.5.4 Discussão do tópico

CETE. A professora, na sua atuação, sustentou a introdução e discussão dos conteúdos enquanto os alunos trabalhavam nas tarefas e, simultaneamente, foi colocando questões e orientando o trabalho a realizar, num ambiente que proporcionou uma participação ativa dos seus alunos nas aulas. Deu, por várias vezes, indicações dos passos a seguir na calculadora gráfica para os alunos alcançarem os resultados pretendidos, de modo a facilitar a adaptação deles à manipulação da calculadora gráfica no tema.

Ao longo das tarefas, a relação bivariada foi caraterizada pela tendência global dos dados, matematicamente expressa pela reta de regressão (também referida como modelo linear). Para não explicar diretamente como se determinava a reta de regressão na calculadora, Estela optou por relacioná-la, de certa forma, com conhecimentos que os alunos tinham recentemente adquirido, na disciplina de Matemática, no tema das Funções, no período anterior. Ou seja, Estela solicitou aos alunos para que procurassem e determinassem a expressão analítica de uma reta que tivesse a particularidade de melhor se aproximar de um conjunto específico de dados bivariados. Só depois introduziu a reta de regressão (obtida da calculadora gráfica) como aquela que melhor se ajustava ao conjunto de dados e, por conseguinte, a mais precisa de todas, sem desejar explicar de uma forma aprofundada o porquê desta situação. A forma como perspetiva o programa escolar pode ter condicionado a sua atuação neste assunto, visto que, nas entrevistas, a professora revelou que a análise da variabilidade dos erros, que poderia ter feito na aula (mas não fez) poderia ter ajudado os alunos a perceber melhor o que estava em causa.

Destaca-se também que o centro de gravidade foi tratado na aula de uma forma marcadamente processual. A interligação entre dados e seu contexto no raciocínio com o modelo ou na explicação de resultados estatísticos nem sempre foi visível.

No estudo de dados bivariados a regressão linear é o único modelo considerado ao nível do secundário; contudo, o averiguar se o modelo linear faz ou não faz sentido através de uma análise à variabilidade dos dados não é considerado pela professora como uma atividade relevante no ensino. Neste tópico, foram realçados muito poucas vezes aspetos importantes do raciocínio estatístico, tais como a consideração da *transnumeração*, atenção à variabilidade e integração de Estatística e contexto.

CA. Estela detetou algumas dificuldades ao longo da resolução das tarefas e tentou ajudar os alunos a superá-las. Uma dessas dificuldades prendeu-se com a utilização da calculadora gráfica. Estela não tinha dúvidas que iriam emergir dificuldades relacionadas com a utilização do *menu* Estatística da calculadora gráfica nas aulas, visto ser a primeira vez que os alunos a usavam para atender à especificidade do tópico.

Estela notou também que alguns alunos não estavam a perceber porque é que a reta que estimaram à mão e a reta de regressão que obtiveram diretamente da calculadora gráfica não coincidiam. Os argumentos que a professora usou nessa situação não estariam a convencer totalmente os alunos; ainda assim, não avançou com outras explicações. A sua ação poderá ser sido influenciada pela sua leitura do programa escolar.

Nas tarefas de associação entre valores de coeficientes de correlação e diagramas de dispersão, houve um acesso limitado aos raciocínios dos alunos e as discussões que ocorreram em torno da análise dos diagramas de dispersão não foram aprofundadas. Na interação em grande grupo, ao discutir a tarefa *População residente em Portugal* (t19), Estela questionou os alunos sobre a utilidade do modelo para se estimar a médio e longo prazo e, ao ouvir a resposta de um aluno, considerou-a "uma explicação puramente matemática"; no entanto, acabou por não questionar o aluno nem assumir que essa resposta não era totalmente adequada. A sua atuação, neste caso, pode ter induzido alguns alunos em erro quanto à aplicabilidade do modelo linear.

CE. A leitura da relação bivariada e da correlação linear no diagrama de dispersão foi intuitivamente associada à reta de regressão. Uma interpretação da reta de regressão que expusesse os dados e o seu contexto não foi evidente na prática de Estela. Quanto ao centro de gravidade, Estela assumiu-o como um ponto pertencente à reta de regressão, ensinou como determiná-lo e representá-lo na calculadora gráfica, mas não se deteve na exploração da sua importância, utilidade e interpretação. Ao mesmo tempo, na entrevista, não revelou entender de uma forma mais profunda a relação do centro de gravidade com a reta de regressão. Estela explicou que a reta de regressão é a reta obtida pelo método dos mínimos quadrados e, embora nas aulas não tenha feito referência explícita a algumas ideias chave relacionadas com a sua composição (desvios, intervalos adequados para se realizar estimativas da variável independente, etc.), apresentou-a como modelo não válido a médio e longo prazo. Reconhecer a fiabilidade do modelo linear foi algo brevemente comentado numa entrevista; ainda assim, foi possível observar que, para Estela, a força do coeficiente de correlação tem um papel preponderante neste reconhecimento. A necessidade de análise à resistência dos coeficientes de correlação em diagramas de dispersão não foi levantada pela professora.

CC. Neste tópico, Estela evidenciou dar importância a conceitos e representações e algumas relações entre eles. A reta de regressão como *a melhor reta que se aproxima dos pontos do diagrama de dispersão* e o coeficiente de correlação linear como medida que indica a força e sinal da relação bivariada foram as ideias mais enfatizadas no tópico, bem como a utilização da calculadora gráfica na obtenção de representações numéricas e gráficas. O ter consciência que poderia ter usado o ecrã da calculadora gráfica com os erros ou desvios representados sobre a reta de regressão, que um aluno alcançou na sua calculadora, pode ter influenciado a sua perceção futura do currículo.

5.3.6 Discussão do tema à luz de cada domínio do conhecimento didático

Conhecimento do ensino do tema de Estatística (CETE). A prática de Estela revela que a resolução de tarefas com o apoio da calculadora gráfica é a atividade principal de cada aula. Este é o contexto privilegiado para introduzir algum conteúdo ou explicar algum assunto, e para solicitar a participação dos alunos, colocando-lhes questões e respondendo às suas dúvidas. A professora seguiu os conteúdos programáticos de acordo com a ordem que estes surgem no programa. Em todos os tópicos, com exceção do primeiro, *População e amostra*, foram propostas nas aulas tarefas diversificadas. Representações gráficas variadas foram surgindo ao longo do ensino do tema, devido ao incentivo que os alunos recebiam da professora para usarem qualquer das calculadoras gráficas que tivessem à mão na resolução das tarefas. Contudo, o estabelecimento de conexões entre representações que se baseavam nos mesmos dados, com vista à promoção do raciocínio estatístico, não foi frequente. Nas análises de dados efetuadas na maioria das situações propostas, as capacidades de *transnumerar*, de analisar a variabilidade para aprofundar algum conceito estatístico ou para interpretar os dados e resultados à luz do seu contexto, de raciocinar com um modelo e de integrar Estatística e contexto tiveram uma presença pouco consistente, ao contrário dos cálculos e das representações, ambos obtidos através da utilização da calculadora gráfica. Para Estela, o tópico das *medidas de dispersão* tinha uma certa ligação com o das *medidas de localização*, embora essas conexões tivessem sido estabelecidas essencialmente ao nível dos cálculos, acabando por não emergirem comentários ou ideias intuitivas que incidissem sobre esses resultados. Estas situações têm repercussões ao nível do desenvolvimento das aprendizagens dos alunos.

Conhecimento do aluno (CA). A maioria das dificuldades detetadas no tema pela professora, na aula, tinha a ver com a utilização da calculadora gráfica na resolução das tarefas propostas. Contudo, poderia ter havido um maior acesso, por parte da professora, aos raciocínios e explicações dos alunos dado que eles eram, de uma maneira geral, bastante atentos e participativos. Também houve dúvidas e comentários de alunos que surgiram na aula e que poderiam ter sido melhor aproveitados pela professora, no sentido do aprofundamento dos assuntos e, por conseguinte, da promoção de uma aprendizagem mais significativa. Tal pode ter acontecido, por um lado, devido a algumas lacunas no conhecimento de Estatística de Estela (por exemplo, relativamente a ideias sobre a representatividade de uma amostra, ou à aplicação da reta de regressão e seu entendimento como um modelo) e, por outro, devido à professora considerar que essas situações tinham a ver com assuntos do programa que não mereciam atenção.

Apesar de Estela assumir que na aprendizagem deste tema foi sobretudo privilegiado um trabalho em torno dos cálculos e dos procedimentos, também ficou um pouco surpreendida pelos erros que encontrou nalgumas respostas escritas de alunos a tarefas que lhes sugeriu para trabalho de casa, tais como no cálculo de estatísticas estudadas e na construção de certas representações trabalhadas em aula. Outro aspeto que surpreendeu a professora foi o facto de a maioria dos alunos ter optado por não explicar por escrito os resultados obtidos, mesmo quando o enunciado da tarefa fazia esse pedido.

Conhecimento de Estatística (CE). A professora revela saber usar a calculadora gráfica e conhecer mais do que uma forma de chegar a certos resultados de estatísticas e representações; além disso, tem a capacidade de sugerir instruções alternativas mesmo para calculadoras com que não tenha tanta experiência. Foram associadas algumas ideias desadequadas a alguns conceitos estatísticos e representações devido a lacunas no seu conhecimento de Estatística e também devido a alguma falta de experiência em lidar e raciocinar com vários assuntos.

Conhecimento do currículo (CC). Para Estela, o programa escolar no tema de Estatística engloba a Estatística descritiva com as suas noções e representações numéricas e gráficas, e não faz explicitamente apelo à interpretação dos resultados, à introspeção dos dados e à realização de inferências. O trabalho conduzido por Estela na aula reflete, de uma maneira geral, as suas ideias sobre o programa escolar neste tema. A sua prática centrou-se bastante na determinação de estatísticas e na obtenção das representações gráficas por meio da calculadora gráfica.

CAPÍTULO 6

O CASO DE LIA

Este capítulo é dedicado ao caso da professora de Matemática Lia. É um capítulo constituído por três partes principais, *Percurso e contexto profissional*, *Perspetiva geral sobre o conhecimento didático do professor em Estatística* e *A Estatística na sala de aula*. Na primeira, é realizada uma breve apresentação da professora, descrito o seu percurso e contexto profissional, destacando-se aspetos marcantes da escolha da sua profissão e da relação que possui com colegas e alunos no contexto escolar. É também dada atenção à sua participação em projetos e formações e à sua experiência na Estatística enquanto aluna e formanda. Adicionalmente, apresenta-se uma breve caraterização da turma onde foi acompanhado o trabalho de Lia.

Na segunda parte apresenta-se uma perspetiva geral do *conhecimento didático em Estatística* de Lia relativamente a alguns aspetos, nomeadamente, à forma como o tema Estatística do programa escolar do secundário é perspetivado pela professora e à leitura que a professora faz dos conteúdos a abordar à medida que percorre os tópicos principais. É ainda caracterizado o modo como Lia habitualmente procede para planificar a unidade de Estatística, destacando os materiais e recursos em que se apoia nessa preparação, e tenta-se compreender o que entende por tarefa estatística. Esta parte é encerrada com uma discussão guiada pelos domínios do modelo do conhecimento didático em Estatística.

Na terceira parte analisa-se o conhecimento didático em Estatística de Lia, tendo por base a sequência de conteúdos estatísticos que percorreu ao longo da sua prática com a sua turma de 10.º ano do secundário, nomeadamente através dos tópicos *População e amostra*; *Organização de dados*; *Medidas de localização*, *Medidas de dispersão* e *Dados bivariados e regressão linear*. Esta parte, além de incluir uma discussão de cada tópico guiada pelos domínios do modelo do conhecimento didático, também incorpora uma discussão, com a mesma estruturação da realizada por tópicos, mas focada no tema Estatística, numa perspetiva geral. Ainda assim, há a referir que os dados dessas discussões apresentam-se mais sistematizados, relacionados e aprofundados no capítulo das conclusões do estudo (capítulo 7), o qual tem a particularidade de integrar uma discussão alargada do conhecimento didático em Estatística das duas professoras participantes do estudo.

6.1 Percurso e contexto profissional

6.1.1 Breve apresentação da professora

Lia é professora de Matemática do 3.º ciclo e do ensino secundário, é casada e tem um filho. Ingressou na faculdade, numa universidade pública do Norte, no curso de Matemática, ramo educacional, no início da década de 90. Deste curso fazia parte um estágio integrado com a duração de um ano letivo, no qual Lia lecionou duas turmas do ensino básico. É docente há mais de 15 anos e já passou por sete escolas distintas afetas à DREN. Já lecionou todos os níveis desde o 7.º ao 12.º ano, mas apenas nas duas últimas escolas lecionou Matemática no ensino secundário.

6.1.2 A escolha da profissão

Lia considera que foi sempre uma boa aluna a Matemática no ensino básico e no ensino secundário. No entanto, foi no ensino secundário que a disciplina de Matemática lhe despertou maior interesse, ao ser lecionada pela mesma professora durante os 11.º e 12.º anos. Lia acha que esta professora lhe apresentou a disciplina de uma forma bastante organizada. Esta situação, juntamente com o facto de gostar de realizar exercícios e de responder a desafios, bem como ser das disciplinas mais práticas do secundário, fez com que a Matemática se tornasse a sua disciplina preferida nesses anos de escolaridade. Contudo, quando concorreu a nível nacional para frequentar a universidade, estava indecisa entre Economia ou Matemática (educacional), tendo acabado por concorrer a ambos os cursos. Na escolha do curso, o critério que mais pesou foi fundamentalmente a proximidade de casa, acabando por escolher frequentar o curso de Matemática numa universidade pública do Norte.

A sua experiência neste curso, desde o primeiro ano, foi bastante positiva, considerando que "tinha encontrado a vocação" daí ter confessado que, desde então, nunca mais pensou em ser economista ou gestora de empresas (entrevista 1). Afirmou também que a sua experiência como professora durante o estágio pedagógico lhe permitiu perceber que estava bem preparada ao nível dos conteúdos a ensinar. Já na parte de relacionamento com os alunos sentiu um desafio muito maior, por ter tido duas turmas bastante complicadas; no entanto, acha que, mesmo nessa altura, conseguiu superar dificuldades e evoluir como professora:

> Aliás, ali [estágio] o que estava mais em causa era o nível de… de pedagogia… de relacionamento com os alunos… Houve algumas críticas que me fizeram no início [pelo orientador] que dizia respeito à [minha] postura na sala de aula… que me ficaram mesmo… Acabei por ser considerada a melhor estagiária [dos 4 estagiários de Matemática da mesma escola] no final. (entrevista 1)

6.1.3 A escola e relação com colegas e alunos

A direção da escola onde Lia atualmente leciona parece considerar importante que os professores tenham uma sala para trabalhar em conjunto. Lia tem, no seu horário, às terças-feiras, tempos letivos para trabalhar com os colegas do mesmo grupo disciplinar. Assim, tem por hábito reunir-se com os colegas de Matemática numa mesma sala, na qual costuma trabalhar essencialmente com os que se encontrem a lecionar os mesmos níveis de escolaridade, nos "chamados trabalhos do subgrupo":

> [Nessa sala]… Planificamos … fazemos as planificações de cada período, planeamos algumas aulas, ou dizemos o que é que realmente poderemos abordar de uma forma ou doutra… partilhamos ou elaboramos materiais, recursos… E fazemos tudo, dentro do grupo, trabalhamos aquilo que achamos que é necessário. (entrevista 1)

Lia também realçou que é sobretudo com os colegas que lecionam os níveis do ensino secundário que mantém um maior contacto dentro e fora da escola, com os quais discute alguns conteúdos a ensinar e fala sobre as situações emergentes do contexto escolar. Referiu ter apenas uma colega de Matemática que também se encontrava a lecionar a uma outra turma de 10.º ano, com quem partilhou alguns materiais (enunciados de fichas de trabalho e de testes) desde o primeiro período.

Lia confessou que gosta da sua profissão. No seu exercício, tem-se deparado com turmas do secundário em que vários alunos vêm do básico com lacunas nos seus conhecimentos de Matemática "mesmo ao nível dos procedimentos e na interpretação das atividades propostas"; aliás, não se lembra de ter tido "uma turma [do secundário] que fosse maioritariamente boa" (entrevista 1). Considera que esta situação pode influenciar o ritmo de trabalho e o ambiente global das aulas. Já por causa disto assume um papel mais participativo no decorrer das aulas, em que procura estar mais próxima desses alunos, para tentar que estejam mais atentos e respondam efetivamente às questões que lhes coloca.

6.1.4 Participação em projetos e formações

Lia mencionou que não tem participado em projetos escolares ou encontros na área do ensino da Matemática, tendo sido a formação sobre Estatística dinamizada pela investigadora a primeira que frequentou sobre este tema. Referiu também que tem frequentado ações na área das tecnologias que são oferecidas gratuitamente por alguns centros de formação do Norte, nomeadamente, sobre a plataforma *moodle* ou quadros interativos.

Considera que essas formações de natureza tecnológica mostram ferramentas que, devidamente ajustadas, podem ser úteis para a sala de aula no desenvolvimento de conteúdos específicos (lembra-se, por exemplo, de ter usado o quadro interativo no ensino básico), além do seu caráter atrativo para os alunos. Afirma que tem usado o Excel no ensino do tema da Estatística nas disciplinas de MACS, como alternativa à calculadora gráfica, até porque os alunos de MACS não são obrigados a possuir calculadora. Lia destaca a utilidade do *Excel* na construção de tabelas e gráficos, e na aplicação dos menus específicos da Estatística:

> … A utilização dada ao *Excel* [com muitos dados], por exemplo, na [formação de] Estatística… tendo tempo… Não sei se vou ter tempo este período, mas, por exemplo, com os módulos dos cursos profissionais, há mais do que tempo para os alunos trabalharem no computador e saberem fazer gráficos, tabelas, saberem

> proceder às técnicas… usarem tudo o que der para usar… E depois tentar que façam uma análise, não é? (entrevista 1)

Para Lia, a tecnologia pode ser particularmente proveitosa na análise de dados. Pondera, se tiver tempo, na sua turma de 10.º ano, utilizar, para além da calculadora gráfica, que é um recurso obrigatório, o Excel, em particular, na construção de tabelas.

6.1.5 A experiência pessoal com a Estatística

Lia não se recorda de ter aprendido Estatística enquanto aluna do ensino básico e secundário. Enquanto aluna do ensino superior acha que só teve uma disciplina na qual aprendeu Probabilidades e Estatística, mas não está certa do seu nome. Do que se recorda é que os assuntos tratados lhe pareciam difíceis por terem sido abordados de uma "forma abstrata" e que sentiu dificuldades semelhantes na disciplina de Física, "… por não ter aquelas bases… Tive de estudar o que estava para trás, para entrar melhor nesses conteúdos" (entrevista 1). Por conseguinte, para lecionar a Estatística, quer do ensino básico, quer do ensino secundário, teve de se preparar. Essa preparação foi feita individualmente e foi focada nos manuais escolares e em pesquisas na internet sobre materiais e tarefas estatísticas. Recorda-se que, da primeira vez que lecionou Estatística, baseou-se fundamentalmente no manual adotado.

Lia comentou ter apreciado, durante a ação de formação, a análise feita aos vários excertos apresentados de interações entre professor e alunos enquanto estes exploravam tarefas de cariz investigativo:

> A parte dos diálogos… no fundo… mais a parte pedagógica em si … Os conteúdos são aqueles… a forma de os explorar é que faz a diferença, acho eu!... Achei interessante a forma como o professor guiava os diálogos… Também a maturidade das respostas dadas pelos alunos era outra… mas isso é algo que temos de desenvolver nos [nossos] alunos… Isto é também algo que gostava de fazer com eles, fazer com que pensem… tirem conclusões. (entrevista 3)

A sua experiência como formanda na ação de formação sobre Estatística dinamizada pela investigadora contribuiu, de certa forma, para que se predispusesse a repensar o papel do professor na orientação das atividades propostas nas aulas e em fazer com que os alunos desenvolvessem gradualmente hábitos de comunicação, passando a retirar conclusões enquanto exploravam as tarefas estatísticas.

6.1.6 A turma e contexto letivo

A turma em que foi analisada a prática de Lia é composta por 25 alunos, 15 do sexo feminino e 10 do sexo masculino. Lia considera que se trata de uma turma em que os alunos são, de uma maneira geral, pouco ativos e pouco participativos, tendo alguns uma postura completamente desinteressada na aula. Também considera que a maioria revela falta de hábitos de estudo, não interligando o que aprendem nas aulas e havendo poucos a realizar o trabalho de casa que estabelece. Em termos das classificações e do desempenho da turma, Lia é da opinião que há "quatro alunos bons, os restantes vão de razoável a fraco" (entrevista 1). Adicionalmente, considera que a turma possui outros quatro alunos que não estão no curso adequado por revelarem muitas dificuldades em Matemática e nas disciplinas que envolvem cálculos e raciocínio matemático.

A escola de Lia é uma escola antiga. As salas de aula estão todas equipadas com mesas individuais agrupadas para dois alunos, um projetor, uma tela e, na mesa da professora, há um computador. As aulas observadas ocorreram em salas similares, exceto a última que foi realizada numa sala de computadores. Nessa sala, havia filas de mesas encostadas à parede e também algumas filas ao centro. Cada uma dessas mesas tinha um computador.

O ambiente nas aulas pode-se caraterizar como informal, com um ritmo de trabalho constantemente imposto pela professora. A relação entre professora e alunos e entre alunos era cordial, sendo estes, de uma maneira geral, disciplinados. No entanto, durante as aulas, a professora chamava-lhes a atenção por serem pouco participativos e por alguns revelarem desinteresse geral pelas atividades, tentando também atenuar a sua tendência para se distraírem com conversas com os colegas de carteira.

Lia é da opinião que os alunos não estranharam a presença da investigadora na sala de aula e que atuaram como habitualmente, confessando que eles até lhe perguntaram quando é que a investigadora iria voltar à sala de aula quando esta não esteve presente na aula de revisões. Só nas duas últimas aulas, alguns alunos chamaram a investigadora ao lugar pedindo ajuda para lidar com a calculadora gráfica. Nas restantes aulas, apesar de não a terem abordado sobre os assuntos da aula, permitiam sempre que esta filmasse as suas resoluções no caderno diário.

A maioria das aulas de Lia foi iniciada com a professora a escrever no computador e a ditar para os alunos, em simultâneo, o sumário da aula. Os momentos de exposição de conteúdos e correção do trabalho de casa ou de alguma tarefa foram conduzidos, fundamentalmente, pela professora, com alguma participação dos alunos. Para incentivar a participação dos alunos na aula, a professora tinha a preocupação de os envolver nas atividades, solicitando-lhes que respondessem às perguntas que lhes colocava e, por vezes, pedindo-lhes para irem ao quadro apresentar a resolução de alguma Tarefa. A maioria dos alunos possuía calculadora gráfica da marca *Casio*. Nas poucas ocasiões em que Lia recorreu à projeção da calculadora gráfica na tela da sala de aula, demonstrou as instruções desejadas nas duas marcas mais usadas pelos alunos, *Casio* e *Texas Instruments*. Este recurso foi usado pontualmente pela professora no ensino do tema de Estatística; contudo, no decurso destas aulas, os alunos frequentemente recorriam à calculadora gráfica sempre que sentiam necessidade.

6.1.7 Síntese

Lia é uma professora com alguma experiência no ensino secundário. É prestável e ativa e mostra-se disponível para colaborar com os colegas da escola e também para participar em formações.

A professora pensa que os seus conhecimentos sobre Estatística foram, sobretudo, adquiridos durante a sua prática profissional, tendo tido necessidade de estudar os diferentes tópicos a lecionar, enquanto preparava as aulas. Os manuais escolares e alguns *sites* de internet apoiaram essa preparação. No ensino da Estatística, dá relevância ao uso das tecnologias. Mais concretamente, na disciplina de Matemática A, usa a calculadora gráfica, porque todos os alunos a possuem, e em MACS recorre a salas de computadores para os alunos poderem usar o *Excel*, uma vez que não são obrigados a adquirir calculadoras gráficas. A sua participação na ação de formação de

Estatística fê-la repensar o modo de interagir com os alunos enquanto eles exploram as tarefas para ajudá-los a desenvolver a capacidade de comunicar resultados.

Na turma onde decorre este estudo, Lia destaca que os seus alunos são, de um modo geral, pouco participativos, contudo, respeitadores. O trabalho desenvolvido nestas aulas é focado em momentos expositivos feitos pela professora e na resolução de tarefas propostas. Sobre a resolução destas tarefas, a professora tem um papel preponderante na sua discussão e correção bem como em envolver os alunos no trabalho proposto.

6.2 Perspetiva geral sobre o conhecimento didático do professor em Estatística

6.2.1 O programa da disciplina de Matemática A

Na leitura que faz do programa escolar, Lia indica que o tema da Estatística do programa de Matemática A do 10.º ano (ME, 2001), com os seus três grandes subtemas (Generalidades, Organização e interpretação de carateres estatísticos e Distribuições bidimensionais), diz essencialmente respeito à "Estatística descritiva" (entrevista 1). Considera que é um tema que se encontra bem organizado, em que os tópicos são adequados e interessantes. Adianta ainda que "a forma como se vai abordar os conteúdos, como se vai apresentá-los aos alunos…" (entrevista 1), bem como o tempo disponível para tratar este tema na aula, podem influenciar a aprendizagem dos alunos. Considera que, para o tema de Estatística nas disciplinas de Matemática dos cursos profissionais do secundário (MACS), que são trabalhadas por módulos, as coisas são diferentes: além de estarem estabelecidos mais tempos letivos para lecionar os conteúdos, os programas são um pouco mais extensos neste tema, dado incluírem alguns aspetos de Estatística inferencial.

Ao refletir globalmente sobre o tema da Estatística no currículo do ensino básico, Lia afirma que, nos 8.º e 9.º anos, são essencialmente desenvolvidas capacidades ao nível de aplicação de conceitos através de cálculos ou leituras elementares de gráficos, acrescentando: "É o que eles sabem resolver… quando têm que pensar mais um bocadinho é que se nota logo que eles não conseguem fazer" (entrevista 1). Relativamente ao ensino secundário, Lia considera que os alunos não saem do secundário devidamente preparados neste tema, dado existir, na sua perspetiva, um fosso grande entre os conhecimentos adquiridos no tema de Estatística da disciplina de Matemática A e na Estatística com que muitos alunos vão ter de lidar no ensino superior, não necessariamente através de disciplinas específicas de Estatística. Refere ainda que:

> Qualquer aluno do superior [de áreas, tais como: Línguas, Saúde, Economia, etc.] tem de saber analisar e interpretar um gráfico, uma tabela referente a um estudo qualquer, de qualquer área (…) No ensino secundário, só há Estatística no 10.º ano… e no 12.º ano tens as Probabilidades e pouco falas da Estatística… só dizes que elas [Estatística, Probabilidades] estão relacionadas uma com a

> outra… mais nada! Por isso é que nós, docentes, damos pouca ênfase à Estatística… É trabalhada a correr… O próprio currículo assim o faz, não é?… Depois, no 12.º ano há um exame, onde nem sequer a Estatística surge e no qual sai Probabilidades… (entrevista 1)

Assim sendo, Lia é da opinião que o tema da Estatística, à semelhança dos outros temas inseridos no programa, deveria ser lecionado ao longo dos três anos do ensino secundário, para ser dada a oportunidade de aprofundamento dos conceitos, e de se poder enveredar por propostas mais desafiantes na aula. Ainda assim, não vê necessidade de uma extensão do programa do ensino secundário, por exemplo, à Estatística inferencial. Para Lia, seria mais importante desenvolver um ensino mais focado na compreensão dos conceitos e das representações sugeridas no programa com maior ligação a situações reais:

> **Professora:** Se eu vou lecionar este tema em três semanas [tempo que considerava expectável], eu não tenho muito tempo para fazer grandes coisas… Portanto, praticamente é só dar os conteúdos e os cálculos… Portanto, se eu tivesse mais tempo para dar, um pouco no décimo, um pouco no décimo primeiro e um pouco no décimo segundo… se calhar… no décimo poderia dar os conteúdos, depois no décimo primeiro poderia revê-los, mas começar a… a dar outra perspetiva da Estatística [fases do processo investigativo]…
> **Investigadora:** Podes exemplificar?…
> **Professora:** Em que… em que os alunos realmente poderiam desenvolver a capacidade de pesquisa, de análise perante qualquer tipo de assunto da realidade, não é?... As capacidades de organizar os dados… de recolher dados… que são… aliás… essas capacidades são pouco desenvolvidas no décimo ano, dado que o tempo é mesmo muito pouco. (…) Portanto, ao fim ao cabo, os conteúdos acabam por ser praticamente os mesmos, também os procedimentos… O que é que se dá [atualmente]?... As tabelas, as medidas de localização, as medidas de… de dispersão, a correlação… Só que muitas das vezes isto é lecionado tão pela rama, que eles não conseguem entender muitas vezes o que é [por exemplo] uma correlação, não é? Porque não têm tempo para interiorizar… para consolidar… para fazer análises mais completas… (entrevista 1)

Lia não tem dúvidas que a Estatística escolar é um tema útil para os alunos, considerando que o acesso à informação estatística está cada vez mais generalizado, por meio de tabelas e gráficos em diferentes meios de comunicação social. A preparação que os alunos obtêm depende muito do tempo letivo que sobra para a Estatística no final do 10.º ano. Considera que os professores dariam, certamente, a devida relevância ao tema se ele fosse incluído nos exames nacionais do secundário. Na sua perspetiva, este tema é considerado fácil para professores e alunos, mas pode constituir um desafio muito maior quando são propostas tarefas cujas questões pedem comparações e discussão de resultados.

Quando questionada sobre a realização de estudos estatísticos que envolvam as quatro fases do processo de investigação estatística, a professora considera que este não é explicitamente referido no programa de Matemática A e, por conseguinte, nunca realizou uma investigação na aula. Contudo, pondera a possibilidade de fazer uma pequena investigação com a sua atual turma, caso tenha tempo. Esta teria de ser feita em grupos de dois ou três alunos e abarcaria "eles escolherem o estudo que queriam fazer… escolher a variável, recolher dados, teriam de fazer análises, não só os gráficos, escrever e retirar conclusões… e apresentarem na aula" (entrevista 1). Também levantou a hipótese de os alunos fazerem um estudo investigativo com variáveis bidimensionais, através de um inquérito e acerca de um tema que lhes chamasse a atenção. Comentou ainda que o momento ideal para ser desenvolvido na aula um trabalho desta natureza seria depois de os alunos aprenderem todos os assuntos do programa, "para terem um pouco mais de autonomia", nomeadamente, no uso da tecnologia para "acederem às representações gráficas" (entrevista 1).

6.2.2 Planificação da unidade de Estatística

Lia considera que o número de aulas que tem para lecionar a Estatística no terceiro período é relevante na definição do grau de profundidade com que trabalha qualquer tema da disciplina de Matemática. Quando inicia qualquer tema ou tópico, tem a preocupação que os primeiros exemplos ou tarefas permitam contemplar uma componente de motivação para o tema. Na sua opinião, a Estatística tem a particularidade de ser um tema que está sempre relacionado com alguma situação da vida real: "Ela relaciona-se facilmente com a vida real… porque há estudos estatísticos seja do que for" (entrevista 1). Lia tem sempre o cuidado de relacionar qualquer tema do programa com situações do quotidiano, sobretudo na sua introdução.

Quer na Estatística, quer em qualquer outro tema, esta professora usa bastante os manuais, sobretudo o adotado, para propor tarefas para a sala de aula e para o trabalho de casa. Também prepara fichas de trabalho nas quais inclui tarefas selecionadas de manuais, da internet e, por vezes, algumas criadas por si própria. Assume, sem complexos, que usa muito os manuais escolares, não deixando de referir que possa haver colegas que tenham relutância em usar o manual adotado na aula, por considerarem que, ao fazê-lo, demonstram, de certo modo, menor capacidade para exercer a profissão:

> Acho que, se calhar, há muito docentes que não querem usar o manual. No sentido de que… Se utiliza o manual, é um professor não tem ideias ou não tem as suas próprias metodologias… Bem… eu acho que é assim: Os manuais existem, não é? São feitos por pessoas que têm competências para tal, não é? Acho que, se os alunos têm o manual… nós temos de o utilizar. Agora, temos de selecionar o que achamos que é interessante ou que é importante… do que é lá proposto, correto? (entrevista 1)

Lia admite que, na altura da primeira entrevista, ainda não tinha tido tempo para analisar com detalhe todo o tema da Estatística no manual adotado. Ainda assim, indicou que situações reais e os conteúdos abrangidos pelo programa lhe pareceram estar contemplados no manual que iria ser usado na aula, bem como nos recursos que utiliza:

> Acho que este manual [adotado]… ainda não o analisei totalmente na Estatística… Foi interessante nos outros temas… Na Estatística, inclui problemas relacionados com situações do quotidiano, com a interdisciplinaridade… Estão lá os conteúdos que têm de ser ensinados… (…) Há um conjunto de recursos que são recursos indispensáveis: o manual, a internet, as fichas de trabalho, o computador para os colocar a fazer tabelas de frequências com o *Excel.* (entrevista 1)

6.2.3 Caraterísticas das tarefas estatísticas

Para Lia, uma tarefa pode ser um exercício, uma atividade que pode ser integrada em fichas de trabalho. Para além das tarefas, as fichas de trabalho podem incorporar definições de conceitos, de representações ou exemplos de situações relacionadas com conceitos a abordar. Na sua prática, propõe fundamentalmente tarefas do manual escolar. Recordou que, nos outros temas matemáticos lecionados no mesmo ano letivo, recorreu pontualmente a fichas de trabalho; já para o tema da Estatística, preparou cinco fichas de trabalho e duas tarefas adicionais (incluídas no anexo 3) com o propósito duplo de lecionar de forma mais rápida e eficiente, dado considerar que acabou por reservar poucos tempos letivos ao tema.

Lia indica que as tarefas que habitualmente propõe na Estatística são as que se encontram nos manuais: tarefas que solicitam a construção de gráficos a partir dos dados fornecidos, e outras que fornecem gráficos e tabelas e em que se tem de dar resposta às questões colocadas sobre estas representações. Especificamente quanto ao uso de dados reais, disse que estes acabam por ser usados na sala de aula porque os "manuais recorrem cada vez mais a dados reais… referem sempre as fontes" (entrevista 1). Lia confessou que nunca pediu aos alunos para recolherem dados e trazerem para a sala de aula, por saber que alguns deles não iriam cumprir o pedido: "o problema é que eles falham muito. Uns trazem, outros não. Nós temos de contar com esses ... imprevistos… Assim vai demorar mais tempo do que planeamos" (entrevista 1). No entanto, já propôs a recolha de dados (qualitativos e quantitativos) na aula, através do fornecimento de informação por cada aluno.

Lia considera que as capacidades que exige ao nível das tarefas trabalhadas na aula de Estatística se centram na construção de gráficos e tabelas e na sua leitura e interpretação, essencialmente, de caráter elementar:

> **Investigadora:** Quais as capacidades que acabam por ser desenvolvidas pelos alunos com as tarefas que propões habitualmente na Estatística?
> **Professora:** Acho que é mais a capacidade de elaborar…
> **Investigadora:** O que é para ti elaborar?
> **Professora:** É… de organizar os dados, fazer gráficos, fazer tabelas, de forma a olharem para a tabela e consigam retirar dali a informação pedida. Há tempo para isso, para a adquirir. Talvez a capacidade de leitura e interpretação de gráficos e tabelas, também … Agora para ir mais além, não sei!…
> **Investigadora:** Por exemplo… pode incluir que expliquem o resultado de uma média ou de um desvio-padrão face ao problema?
> **Professora:** Teriam dificuldades. Este tipo de interpretação… Interpretar uma média, mediana… era preciso mais tempo para ser praticado… para ser desenvolvido. O tempo que se tem não o permite. (entrevista 1)

Quanto ao tipo de questões a colocar nas tarefas de Estatística, Lia também explica que o colocar perguntas de interpretação que tenham a ver com uma determinada situação problemática, à semelhança do que viu na ação de formação, não é habitual e, por isso, acha que os seus alunos teriam dificuldades em responder a questões deste tipo. Na perspetiva de Lia, depois de os alunos aprenderem como aplicar os conceitos e procedimentos de um assunto terão facilidade em resolver qualquer tarefa que especifique exatamente o pretendido:

> Se a tarefa for realmente do tipo 'calcula, determina, faz', eles, já sabendo como é que aquilo se faz, fazem… e não vão ter grandes dificuldades. Mas, se for para interpretar os resultados de uma coisa, saber o que é que acham de uma coisa ou de outra, relativamente a algum assunto, a um estudo realizado sobre uma turma ou em duas turmas, eles aí sentiriam mesmo dificuldades em interpretar o gráfico, o resultado… É o que eu acho… (entrevista 1)

Quando a investigadora apresentou à professora, durante a primeira entrevista, duas tarefas, *Média sobre média* e *Tempo de espera pelo autocarro* (ambas do anexo 2), e a questionou sobre a possibilidade de serem propostas na sala de aula, Lia comentou que os alunos não teriam dificuldades no cálculo da média na primeira; já em relação à segunda, indicou que os alunos só a saberiam resolver se lhes fosse relembrado o cálculo da média para dados agrupados. Ainda na análise desta segunda tarefa, a professora também transmitiu a opinião que os alunos não teriam dificuldades na leitura do diagrama de barras por lhes ser familiar; quanto ao determinarem o desvio-padrão, disse que, depois de lhes ensinar a fórmula ou como chegar ao seu valor na calculadora gráfica, estes alcançariam esse resultado. Na sua opinião, a parte mais difícil desta tarefa, para os alunos, seria "interpretar o resultado do desvio-padrão, acho que era o mais complicado para eles" (entrevista 1). Lia comentou que esse pedido de interpretação não é usual aparecer nas tarefas e também não o costuma fazer.

Lia revela ter noção do maior grau de desafio para o professor que as tarefas exploratórias com vários dados reais podem implicar em termos de orientação aos alunos, bem como nas respostas a dar às questões dos alunos. Considera também que os seus alunos não têm maturidade suficiente para efetuarem, por si mesmos, uma investigação na área da Estatística sem serem acautelados com "orientações bem específicas da parte da docente" (entrevista 1).

Quando a investigadora lhe mostrou a tarefa *Como está a evoluir a nossa população?* (anexo 2), Lia achou que esta seria muito difícil, para a maioria dos seus atuais alunos de 10.º ano, saberem o que fazer: "Este tipo de pergunta [referindo-se à parte das seis orientações]… a linguagem… não está patente nos manuais… eu acho! Teria de estar a aula toda a receber uma série de perguntas: 'o que é para fazer ou o que não é para fazer?'" (entrevista 1). Considera que esta última tarefa "não é direcionada, não é como as outras duas [tarefas] em que já se sabe o que se está a perguntar... A primeira coisa que estão habituados a fazer é organizar os dados em tabelas, em percentagens…"

(entrevista 1). Ao equacionar a hipótese de levar esta tarefa à sala de aula, Lia referiu que não alteraria o enunciado mas que teria o cuidado de explicar cada item pormenorizadamente e de conduzir a sua resolução até a fim. Se a sua resolução corresse bem na turma, talvez propusesse "uma outra semelhante… para ser trabalhada autonomamente" (entrevista 1), de modo a poder ficar com uma melhor perceção das dificuldades dos alunos.

Para Lia, o tema da Estatística pode ser atrativo para os alunos, se o professor procurar assuntos que lhes despertem a atenção. Pela sua experiência na lecionação do tema, não recorda ter-se deparado com dificuldades, quer no ensino, quer na aprendizagem, talvez por se ter centrado nos cálculos e nas representações, devido a constrangimentos de tempo. Contudo, considera que se, neste ano letivo tiver tempo e puder arranjar atividades que fomentem mais comunicação e discussão entre os alunos, à semelhança do que observou/analisou na ação de formação, terá um desafio maior na condução da aula e ganhará, certamente, uma maior consciência das suas dificuldades e das dos alunos:

> Acho que, se eu realmente enveredar por outro caminho, por este tipo de tarefa… investigativa… terei de enveredar por outro tipo de diálogo… como se fez na tarefa que se preparou para a ação [de formação com dados reais sobre o turismo em Portugal]… Na aula foi feita uma abordagem diferente. Colocámos os alunos a dialogar uns com os outros. Nem houve grande confusão porque não estavam todos… apenas uns 13 ou 14 alunos (alguns alunos tinham faltado a essa aula). (…) Penso que irei sentir algumas dificuldades nesta abordagem, só saberei quais depois de experimentar mais. (entrevista 1)

Neste excerto, relembra também a experiência de ter proposto uma tarefa que construiu com uma colega da formação, na sua turma de 10.º ano, que incluía questões de interpretação de um gráfico e outras que pediam alguma discussão à volta de uma situação real. Para Lia, essa experiência era diferente das que habitualmente sugeriu, por estar a propor uma tarefa com questões mais desafiantes para os alunos, que envolvia dar-lhes tempo para analisarem os dados e discutirem as respostas com os colegas e uma maior interação na aula.

6.2.4 Discussão sumária dos dados à luz dos domínios do conhecimento didático

CETE. Lia recorre com bastante frequência ao manual adotado para propor tarefas para as aulas e para o trabalho de casa e, pontualmente, a fichas de trabalho. Considera que as tarefas que sugere na sua prática incluem dados reais e são habitualmente centradas na aplicação dos conceitos aprendidos, nos cálculos e na construção de representações, auxiliando-se, sobretudo, em questões objetivas, que raramente incidem na interpretação de resultados ou em inferências sobre algum assunto.

Lia pondera a hipótese, caso tenha tempos letivos suficientes, de levar para as aulas atividades que solicitem mais interpretação e comunicação dos assuntos estudados, à semelhança do que observou e analisou na ação de formação, ou ainda de propor uma pequena investigação estatística para ser realizada em grupo, colocando os alunos a decidir o tema, recolher os dados, realizar análises com apoio da calculadora gráfica, escrever os resultados e ainda a apresentar o trabalho à turma. Não tem dúvidas que propostas desta natureza são mais desafiantes para os professores e também para os alunos que não estão habituados a comentar, discutir e fazer comparações de conjuntos de dados, exigindo uma orientação constante do professor. Assim, considera que as atividades investigativas deveriam ser propostas depois de abordados todos os assuntos e de dominar-se a calculadora gráfica.

CA. A professora considera que os alunos de uma maneira geral adquirem a capacidade de construir um gráfico ou de fazer um cálculo. Apesar de considerar que os seus alunos não têm dificuldades na aprendizagem do tema, acha que os alunos de uma maneira geral não estão preparados para responder a questões que envolvam interpretação de resultados. O contexto da situação problemática, variabilidade dos dados são alguns aspetos que Lia indicou como desconhecidos para os alunos, ou seja, são aspetos que não fazem parte da linguagem que se usa na aula de Estatística.

Lia, tem também a perspetiva de que os alunos que terminam o ensino secundário não estão devidamente preparados para fazerem leituras a gráficos e tabelas que irão facilmente encontrar no prosseguimento de estudos que venham a fazer.

CE. Lia destaca que todos os tópicos do tema da Estatística da Matemática A são, de uma maneira geral, acessíveis. Como não tem experiência com atividades investigativas, não sabe se terá necessidade de aprofundar alguns conhecimentos estatísticos para melhor lidar com essas situações na aula, caso o programa escolar venha a incluir essas atividades no tema da Estatística.

CC. Lia revela ter uma perspetiva de que o tema da Estatística do programa de Matemática A de 10.º ano é adequado e interessante, foca-se na Estatística descritiva. A professora considera que na prática este tema acaba por ser sobretudo desenvolvido à volta dos cálculos e procedimentos devido a ser um tópico que é lecionado nas últimas aulas do ano letivo. Caso tivesse mais tempo letivo para o lecionar procuraria aprofundar os conceitos e representações, também através de análises de dados que envolvam situações reais. Lia é da opinião de que essas noções e representações deveriam ser trabalhadas da forma mais profunda possível, nomeadamente, ao longo dos três anos de secundário à semelhança dos outros temas lecionados na Matemática. Pelo que sendo a Estatística circunscrita ao 10.º ano, acaba por ser o tema que os professores optam por lecionar de uma forma mais expedita, visto também não ser um assunto coberto nos exames que os alunos realizam no ensino secundário. Para Lia, o programa de Matemática A, não contempla a Estatística inferencial nem as investigações estatísticas, comentando que nunca propôs uma atividade de natureza inferencial na aula.

6.3 A Estatística na sala de aula

6.3.1 População e amostra

In the supermarket, if the product is truly homogeneous, then a small sample will serve the purpose of representing the product. In statistics the issue of sample size is important in balancing economy of scale with the necessity to describe variation in the population as well as possible.
Jane Watson (2006, p. 28)

Este tópico é analisado em três secções. A primeira (noções e dados estatísticos) incide sobre algumas noções centrais e situações apresentadas nas aulas; é explicado como surgiram essas noções e é analisado como foram desenvolvidas na prática de Lia. Esta secção também inclui o tipo de amostras utilizadas no tópico e ao longo do tema. A segunda secção (reflexão sobre o tópico na aula) inclui algumas reflexões de Lia sobre o ensino e aprendizagem do tópico e a terceira secção (discussão do tópico) apresenta uma discussão do tópico guiada pelas quatro dimensões do modelo do conhecimento didático do professor em Estatística.

6.3.1.1 Noções e dados estatísticos

Noções

Lia optou por iniciar este tópico com a introdução de alguns conceitos e exemplos, através de momentos expositivos, antes de propor tarefas para serem trabalhadas nas aulas. A definição de população apresentada pela professora foi a seguinte: “é o conjunto de todos os elementos que apresentam uma ou mais caraterísticas em comum e que se pretende analisar. Os elementos da população podem ser pessoas, animais, instituições, objetos, acontecimentos” (observação 1). A amostra foi definida como “uma parte ou subconjunto da população” (observação 1). Lia associou amostra ao conceito de sondagem quando referiu que uma sondagem dizia respeito à realização de um estudo estatístico sobre uma determinada população “através de uma amostra”. Também referiu a importância das amostras serem representativas e aleatórias e que o recurso à sua estratificação poderia ser útil para o fenómeno que se estava a estudar. Lia indicou especificamente "critérios de seleção de uma amostra" (observação 1):

> **Ampla**: conter um número alargado de elementos, para que os resultados obtidos possam ser generalizados;
> **Estratificada**: estar dividida em subconjuntos ou estratos. Cada estrato deve incluir elementos de forma proporcional à população;
> **Aleatória**: dentro de cada estrato, os elementos são escolhidos ao acaso;
> **Representativa**: conter elementos de todos os estratos da população.

Figura 46: Excerto retirado de um *powerpoint* apresentado na aula (observação 1).

As conexões entre estes conceitos não foram evidentes na interação que houve na sala de aula. O critério "ampla" (figura 46) parece chamar a atenção para o tamanho da amostra, e sua representatividade e aleatoriedade, de forma a possibilitar "os resultados poderem ser generalizados" (observação 1). Os critérios "aleatória" e "representativa" (figura 46) foram explicados tendo como referencial uma amostra estratificada. Lia usou a situação *Qual a amostra mais fiável?* (ver excerto na figura 47) como exemplo de aplicação dos critérios que indicou aos alunos. Mais concretamente, foram lidas as informações fornecidas das três amostras contempladas na situação apresentada e depois a professora questionou os alunos sobre qual das três seria a amostra mais fiável. Vários alunos indicaram que a empresa *Tests* era a melhor. Apesar de Lia ter indicado que esta resposta era a correta, não teceu nenhum comentário, nem questionou os alunos sobre os motivos dessa escolha.

A única diferença entre *Tests* e a *Kmercs* era, efetivamente, o tamanho da amostra: 1000 e 100, respetivamente. A única diferença entre a empresa escolhida pelos alunos da turma (*Tests*) e a terceira empresa envolvida *Marks* recaía na escolha dos participantes (na *Tests* eles foram escolhidos de uma lista telefónica abrangendo diferentes regiões do país, na *Marks* foram escolhidos por terem participado telefonicamente no programa televisivo A). Assim, a amostra da empresa *Marks* é enviesada, mas tal não foi assinalado.

Foram apresentados os seguintes resultados:

Empresa Marks	
Estação A	80%
Estação B	1%
Estação C	7%
Estação D	12%

Ficha técnica: Esta sondagem foi realizada para a estação televisiva A, pela empresa Marks [...]. Foram feitas 1000 entrevistas telefónicas efectivas, das quais 500 a mulheres. Os números de telefone foram escolhidos aleatoriamente de entre os números das chamadas recebidas durante um programa televisivo da estação A.

Empresa Tests	
Estação A	31%
Estação B	15%
Estação C	26%
Estação D	28%

Ficha técnica: Esta sondagem foi realizada para a estação televisiva A, pela empresa Tests [...]. Foram feitas 1000 entrevistas telefónicas efectivas, das quais 650 a mulheres. Os números de telefone foram escolhidos aleatoriamente de uma lista telefónica abrangendo diferentes regiões do país.

Empresa Kmercs	
Estação A	30%
Estação B	13%
Estação C	19%
Estação D	38%

Ficha técnica: Esta sondagem foi realizada para a estação televisiva A, pela empresa Kmercs [...]. Foram feitas 100 entrevistas telefónicas efectivas, das quais 60 a mulheres. Os números de telefone foram escolhidos aleatoriamente de uma lista telefónica abrangendo diferentes regiões do país.

Figura 47: Situação *Qual a amostra mais fiável?*

Lia referiu também, na aula, alguns processos de amostragem, nomeadamente amostragem aleatória simples e estratificada e ainda amostragem sistemática, através da situação *Como reduzir a taxa de absentismo na empresa?* (figuras 48 e 49), que incluía essas opções de amostragem. Lia leu os seguintes excertos:

Uma hipótese é escolher os 20 funcionários sem obedecer a qualquer condição. Para isso, basta atribuir a cada funcionário um número de 1 a 200 e recorrer a um *software* para gerar aleatoriamente, vinte números (funcionários). Podemos recorrer, por exemplo, à calculadora gráfica. → **amostragem aleatória simples**

Outra hipótese é atribuir também a cada funcionário um número de 1 a 200 e a escolha ser feita seleccionando, por exemplo, o 9.º e a partir daí o 19.º, o 29.º, o 39.º, ... e assim sucessivamente, até encontrar os vinte números acabados em 9 que existem de 1 a 200. → **amostragem sistemática**

Figura 48: Excertos da situação *Como reduzir a taxa de absentismo na empresa?*

Ainda outra hipótese será dividir por estratos os funcionários da empresa. Por exemplo, por classes etárias. → **amostragem estratificada**

Determina-se, proporcionalmente, a representação de cada uma das classes etárias e obtém-se o número desejado de pessoas a seleccionar:

Classe etária	Número de funcionários	Percentagem	Seleccionados
Menos de 30	50	25% (= 50/200)	25% × 20 = 5
Entre 30 e 50	140	70% (= 140/200)	70% × 20 = 14
Mais de 50	10	5% (= 10/200)	5% × 20 = 1
Total	200	100%	20

Serão então seleccionados aleatoriamente 5 funcionários com menos de 30 anos, 14 com idades entre 30 e 50 anos e 1 com mais de 50 anos.

Figura 49: Excerto da situação *Como reduzir a taxa de absentismo na empresa?*

Ao revisitar os conceitos de população, sondagem e amostra, Lia teve a seguinte interação na aula:

> **Professora:** Uma população pode ser um conjunto de pessoas, objetos, etc… Ou seja, é tudo o que nós estudamos. (…) Eva, o que é uma sondagem?
> **Eva:** É um estudo estatístico numa população.
> **Professora:** Numa população?
> **Alguns alunos:** Não, numa amostra.
> **Professora:** Ah… É numa amostra… Ou seja, o estudo incide numa porção da população que nós chamamos amostra. (…) Quando é que utilizamos uma amostra?
> **André:** Quando quero fazer uma sondagem…
> **Professora:** Sim, mas… Quando é benéfico utilizar uma amostra? É sempre preciso utilizar a população toda? ... Porque é que, por vezes, escolhemos uma amostra e não a população?
> **Rui:** A população pode ser muito grande…
> **Professora:** E pode demorar muito tempo… E pode ser… moroso e… dispendioso… Então o que é que fazemos?… Temos de selecionar uma amostra! Pode ser selecionada de qualquer forma?
> **Alguns alunos:** Não.
> **Professora:** Se ela fosse selecionada de qualquer forma, o que iria acontecer?
> **Rui:** Os resultados iriam variar.
> **Professora:** Os resultados iriam variar… Quando a selecionamos… Quais as suas caraterísticas?... [Digam] assim mais ou menos…
> **Eva:** Ampla.
> **Professora:** O que quer dizer ampla?
> **Eva:** Abrangente.
> **Professora:** [Digam] Mais?...
> **Nuno:** Aleatória.
> **Professora:** Mais?
> **Martim:** Estratificada.
> **Professora:** Mais?...
> **Hélder:** Representativa da população.

> **Professora:** [De facto] A amostra não pode ser selecionada de qualquer forma porque senão os resultados podiam quê? Podiam sair errados!... (observação 1)

Nesta interação, surgiram algumas ideias que evidenciam o benefício de se trabalhar com uma amostra em vez de uma população. Rui apontou a questão do tamanho e a professora acrescentou tempo e custos, como fatores a ponderar. A professora tentou também destacar que a recolha de uma amostra não podia ser feita de qualquer forma e o aluno Rui referiu que, se esse processo fosse feito de qualquer maneira, "os resultados iriam variar". Lia pareceu concordar com a resposta deste aluno; contudo, não o inquiriu sobre o que queria dizer. Não é possível saber se ele estaria a referir-se a várias amostras ou à chegada de resultados diferentes dependendo do processo de recolha escolhido.

Ainda nesta interação, quando a professora abordou os possíveis critérios de seleção de uma amostra, houve alunos que apontaram os quatro aspetos que a professora tinha previamente contemplado na aula: ampla, aleatória, estratificada e representativa. Contudo, ao explicar o que entendia por ampla, a aluna respondeu apenas "abrangente". Lia não lhe pediu mais explicações nem questionou os restantes alunos acerca do que entendiam com os outros critérios. A professora apelou à necessidade de se usar esses critérios na seleção de uma amostra para não se correr o risco de se trabalhar com uma amostra que produza resultados não fidedignos.

Depois destes momentos expositivos, Lia propôs na aula as tarefas *Intenções de voto em 2006* (t1); *Alunos inscritos no ensino* (t2) e *Serão boas amostras?* (t3). Procedeu à correção de cada uma delas em grande grupo, depois de ter dado algum tempo aos alunos para trabalharem nelas, intercalando momentos de resolução com os de correção.

Tarefa *Intenções de voto em 2006* (t1)

Esta tarefa, proposta na aula, inclui um excerto de uma notícia, um gráfico de linhas, uma ficha técnica, que dá algumas informações sobre a forma como se processou a recolha de dados, e, ainda, cinco questões para serem respondidas pelos alunos. Destas questões, apenas uma requeria a leitura do gráfico; as quatro restantes pediam informações sobre a amostra, que poderiam ser obtidas através da leitura da ficha técnica. Mostra-se, de seguida, a interação que houve entre professora e alunos na correção desta tarefa:

> **Professora:** Marisa, lê a alínea a)… O que é que está a pedir?
> **Marisa:** É uma sondagem.
> **Professora:** Porque é uma sondagem? Explica lá!
> **Marisa:** Porque…
> **Professora:** Quem quer ajudar a Marisa? Diz lá, Sofia.
> **Sofia:** É uma sondagem porque engloba uma parte da população.
> **Professora:** É o que lês na notícia [que foi feita uma sondagem]?
> **Sofia:** [ouve-se a aluna a ler] 'Foram feitas 550 entrevistas telefónicas efetivas'.
> **Professora:** Foram feitas 550 entrevistas. Portanto [essas entrevistas] não foram feitas à população toda. A amostra é de 550. Com isso, já se responde praticamente à alínea seguinte [referindo-se à dimensão da amostra] (…) [Então] De que forma é que ela foi feita [a amostra]?
> **André:** Por via telefónica.
> **Professora:** Muito bem, André. [Agora, a alínea seguinte pede] Quantos homens foram entrevistados?
> **Luzia:** 255 homens.
> **Professora:** Mas eu quero em percentagem.
> **Luzia:** [faz os cálculos no lugar e diz] 46,4%.
> **Professora:** Quantas pessoas correspondem aos 77,8% da taxa de resposta? Quem teve dificuldade em fazer esta questão?... Quem não percebeu a pergunta? … Rute, como fizeste? (professora dirige-se ao quadro para assentar o cálculo que a aluna dita): 550 está para 100% e *x* está para 77,8%. Arredondado às unidades dá?
> **Rute:** 428 pessoas. (observação 1)

A partir do momento que uma aluna diz o que entende por sondagem, a interação da professora tem em vista a identificação da amostra na notícia e, ainda, que os alunos retirem da ficha técnica os valores que lhes são pedidos na tarefa, ora em percentagem, ora em termos absolutos.

Tarefa *Alunos inscritos no ensino* (t2)

Esta tarefa inclui uma notícia retirada da revista *Visão*, de 2006, e um gráfico que indica as percentagens dos alunos do ensino secundário de vários países que frequentam o ensino geral e o profissional. Apresenta-se um excerto da interação da professora com os alunos quando discute com eles a adequação e vantagens e desvantagens de se realizar uma sondagem ou um recenseamento "quando o Ministério da Educação pretende conhecer o abandono escolar por parte dos alunos inscritos em cursos profissionais":

> **Professora:** Que vos parece ser mais adequado? Vocês faziam uma sondagem ou um recenseamento?... O que é uma sondagem? E o que é um recenseamento?... Quero saber isso tudo… Neste caso, trata-se de uma sondagem ou recenseamento aos alunos…
> **Mário:** Sondagem abrange só uma parte da população…
> **Professora:** Abrange só uma parte dos alunos, não é!? … E um recenseamento abrange todos… Se calhar, preferia um recenseamento, diz o André… Qualquer um destes estudos é viável, não é? Só que o recenseamento demora mais tempo, certo? Só que se calhar, para sabermos mesmo a opinião… íamos mesmo para um recenseamento… Mas, há vantagens e desvantagens… É o que diz na alínea seguinte… Em cada tipo de estudo… Num recenseamento, qual será, por exemplo, uma vantagem?
> **Alguns alunos** (falam ao mesmo tempo, ouve-se) Mais resultados… Valores mais concretos…
> **Professora:** Querem dizer… obter resultados mais fiáveis, certo? [No recenseamento] Já sabemos o que dali vamos obter. E qual será uma desvantagem?
> **Alguns alunos:** Demorar mais tempo.
> **Professora:** Demorar muito mais tempo e, se calhar, mais custos. E uma vantagem para a sondagem? Demora menos tempo, portanto, mais rapidez… dá menos trabalho…
> **Eva:** E menos resultados…
> **Professora:** E a desvantagem?... Ah, isso já é uma desvantagem, Eva. Em vez de se obter um valor exato, [obtém-se] um valor aproximado daquilo que se poderá… poderá querer, não é? (observação 1)

Neste excerto, o contexto da situação problemática que deu origem à interação acabou por não ser referido e a discussão foi centrada fundamentalmente nas vantagens e desvantagens de se fazer um recenseamento ou uma sondagem, em termos gerais. Lia acabou por se referir aos resultados de um recenseamento como exatos e fiáveis e aos de uma sondagem como valores aproximados. A professora avaliou as respostas dos alunos quanto a haver mais resultados num recenseamento e menos resultados numa sondagem apenas na perspetiva da qualidade dos resultados. Ou seja, ela não pareceu ponderar a possibilidade de os alunos se estarem a referir à quantidade de resultados que podem ser obtidos de cada tipo de estudo estatístico.

Tarefa *Serão boas amostras?* (t3)

Esta tarefa (t3) expõe, de uma forma sucinta, os processos através dos quais duas amostras iriam ser recolhidas. Foi corrigida em grande grupo, depois de Lia ter dado algum tempo para os alunos analisarem as situações:

> **Professora:** Vamos pensar se estas amostras serão boas amostras ou não.

> (manda o aluno Rui ler a primeira alínea e explicar se a situação corresponde ou não a uma boa amostra)
> **Rui:** (indica o erro) Os utilizadores são todos os estudantes…
> **Professora:** Ah! A maior parte será estudantes… Então não será uma boa amostra… E se escolher uma amostra de dez benfiquistas para se prever o resultado do próximo jogo Porto-Benfica (ouve-se vários risos)? Tu és benfiquista!… (apontando para um aluno). (observação 1)

As duas situações indicadas na tarefa foram declinadas como boas amostras, apesar de, no momento de interação transcrito acima, não terem emergido comentários ou questões nem da parte da professora, nem dos alunos, nem terem sido referidos os critérios mencionados na aula. Pensa-se que a primeira situação terá sido rejeitada como boa amostra por não ser representativa da população (dado ser recolhida à porta de uma escola, seria natural que os participantes fossem, na sua maioria, alunos). A segunda situação provavelmente terá sido entendida como uma amostra enviesada, também por não ser representativa. No entanto, não foram referidos processos de recolha ou amostras alternativas para cada situação, nem se falou da necessidade de estas amostras serem aleatórias.

Dados estatísticos

Os dados estatísticos das tarefas t1 e t2 foram retirados de meios de comunicação social portugueses. Já a tarefa *Será uma amostra?* (t3) não inclui dados mas tem implícita a identificação da amostra e leva a que se levantem questões à sua representatividade. A maioria das tarefas propostas nas aulas de Estatística de Lia incluía conjuntos de dados quantitativos que eram tomados como amostras. Esses dados quantitativos estavam relacionados com variáveis quantitativas discretas ou contínuas. As primeiras foram designadas como aquelas "que assumem um número inteiro positivo"; quanto às segundas, foi dito aos alunos que "podem representá-las com números que podem ir até às décimas, centésimas… é contínua!" (observação 1). Esses dados quantitativos, associados a variáveis contínuas, surgiram agrupados em classes nalgumas tarefas. Além disso, os dados quantitativos das diferentes tarefas, de uma maneira geral, faziam alusão a situações da vida real, embora as tarefas não informassem sobre as respetivas fontes.

Na aula, os alunos não tiveram que proceder à recolha de amostras. O entendimento de que há várias fontes de variabilidade nos dados não surgiu na aula de Lia. Em particular, não foi explorado que, na recolha de dados, essa variabilidade pode ser causada pela imprecisão do instrumento de medição ou por algum erro no registo de informação.

6.3.1.2 Reflexões sobre o tópico na aula

Lia refere que o propósito das tarefas que propôs na sala de aula sobre população e amostra é levar o aluno a saber identificar e distinguir alguns conceitos (população e amostra, sondagem/recenseamento, etc.), entender a característica que está a ser analisada, efetuar alguma leitura de gráficos e alguns cálculos (percentagens), e adquirir a noção que a recolha de amostras é feita adotando-se alguns critérios:

> As tarefas que propus nesta parte era[m] para ver se conseguiam identificar os conceitos básicos. E olhar para o gráfico e saber ler e ver se tiravam de lá informação. (…) [Estas tarefas] Falam mais sobre a população, amostra… o que foi inquirido, o que é que não foi… se foi feito uma sondagem ou recenseamento, calcular percentagens, verificar o tipo de amostras que são feitas, como foram feitas, se foram bem feitas ou não. (…) As tarefas baseadas em estudos que são feitos na comunicação social… não são feitos numa população, são feitos com uma amostra… Foram propostos para eles sentirem que a amostra não é feita de qualquer forma… para que o resultado seja mais próximo da realidade. (entrevista 2)

Lia referiu que, embora não tivesse a intenção de aprofundar o processo de amostragem, considerou mostrar exemplos em que foram usados processos diferentes e válidos, “pois a amostragem não pode ser feita de qualquer forma” (entrevista 2). Quando questionada sobre a razão pela qual referiu em sala de aula quatro critérios (ampla, estratificada, aleatória e representativa) para a escolha de uma amostra, sendo que apenas dois deles estavam explicitamente referenciados no manual (representatividade e aleatoriedade), Lia explicou:

> **Professora:** No fundo… senti necessidade de detalhar mais a informação nessa parte… para lhes mostrar como é que poderia ser a amostra… Ao fim e ao cabo… as três primeiras [ampla, estratificada, aleatória] estão incluídas na representatividade… Ser estratificada está relacionado [com a representatividade], dividida em subconjuntos e cada subconjunto ou substrato dever incluir elementos de forma proporcional à população… Depende do estudo que se quer fazer, não é?… Ampla… quanto mais ampla for… maior o número… mais se vai aproximar do resultado… Portanto, se calhar no manual não estava tão… uso vários manuais… Vi num outro manual qualquer essas

> quatro… achei que eles [alunos] assim entendiam melhor para não ser só insinuado…
> **Investigadora:** Usaste exemplos…
> **Professora:** Neste exemplo que dei [Qual a amostra mais fiável?], os alunos viram logo… que a amostra de 100 não era ampla, não é?... Eles disseram que a de 1000 era melhor…[mas como havia duas com 1000] eles escolheram a que tinha sido feita por estratos e [nessa opção] os números de telefone foram escolhidos aleatoriamente abrangendo diferentes regiões do país; cá está, estratificada, … Ali já não… Essa amostra aleatória foi escolhida dentro de um programa que se estava a ver que era totalmente diferente… eles [alunos] perceberam logo qual era a melhor…
> **Investigadora:** Lia, a amostra representativa tem de ser sempre estratificada?
> **Professora:** Acho que… depende do que se está a estudar, não é?… Mas, por exemplo, neste caso [Como reduzir a taxa de absentismo na empresa?], se fosse escolher uma amostra só dentro dos que têm menos de 30 anos… não era representativo, só estaria a escolher pessoas de uma faixa etária… Depois… até pode calhar bem, mas também pode calhar muito mal… os resultados. (entrevista 2)

Lia deixa transparecer que considera que ser ampla, aleatória e estratificada são caraterísticas importantes de uma amostra representativa. Entre duas amostras aleatórias de tamanhos iguais (1000) feitas a um grupo específico de pessoas (eventualmente enviesadas) ou feitas por estratos, prefere a última. Entre duas amostras aleatórias de tamanhos (100 e 1000) diferentes e retiradas com recurso à estratificação, prefere a de maior tamanho. Deste modo, parece revelar que os estratos podem servir para capturar melhor caraterísticas da população e permitem aceder a uma amostra mais representativa da população, assim como as amostras de maior tamanho.

6.3.1.3 Discussão do tópico

CETE. Lia recorreu a alguns momentos expositivos para apresentar conceitos e algumas caracterizações de amostra e de amostragem, com a ajuda de exemplos. Estas situações serviram, sobretudo, para que os alunos ficassem com alguma perceção da existência de métodos de recolha de dados. Nas tarefas trabalhadas na aula, não foi colocado em prática nenhum dos métodos de amostragem referidos nos exemplos. As três tarefas propostas tinham o objetivo de ajudar os alunos a distinguir e usar os conceitos envolvidos, embora de uma forma superficial. Na aula, na correção das tarefas, Lia deixou transparecer a preferência por recenseamentos em oposição às sondagens, por motivo de serem baseados na população e permitirem alcançar resultados mais fiáveis. A necessidade de recolha de amostras e de se controlar a variabilidade das amostras usadas na aula não foi visível na prática de Lia. O CETE de Lia deixa transparecer a sua visão do programa escolar relativamente a este tópico.

CA. Na aula de Lia houve um acesso limitado aos raciocínios dos alunos. Lia, por vezes, não os questionou sobre as suas respostas e os alunos também não colocaram dúvidas. Lia pode ter induzido os alunos em erro ao ter assumido a sua preferência por recenseamentos comparativamente às sondagens. Por exemplo, os alunos podem ter ficado a pensar que os resultados das sondagens têm pouca credibilidade, quando a amostragem aleatória é um dos métodos mais eficientes, que dão resultados bastante fidedignos e, muitas vezes, com margens de erro pequenas.

CE. A professora parece não conhecer exemplos de estudos que, na prática, só possam ser feitos através de sondagens/amostragens aleatórias visto ter dado mais importância aos resultados provenientes de recenseamentos do que aos das sondagens. Lia sabe que as amostras devem ser aleatórias e representativas e, na aula, abordou mais dois critérios, ampla e estratificada (em relação aos que habitualmente aparecem nos manuais), com vista à sua seleção. Embora, na aula, não surgissem explicações que relacionassem esses critérios, na entrevista, Lia teceu alguns comentários nesse sentido. Os dados sugerem que algumas lacunas ao nível do conhecimento de Estatística da professora neste tópico podem ter influenciado negativamente o conhecimento do aluno.

CC. Para Lia, os assuntos do programa de Matemática A relacionados com população e amostra devem ser mencionados de um modo bastante superficial. No entanto, dá algum destaque aos conceitos que aborda e tem o cuidado de incluir situações da vida real com notícias e gráficos nas tarefas. O seu conhecimento curricular parece estar em consonância com o seu CETE.

6.3.2 Organização e interpretação de dados

Mesmo o gráfico mais simples pode induzir em erro, se não soubermos avaliar qual a informação que ele pretende transmitir!
Graça Martins (2012)

Este tópico é analisado em três secções. Na primeira (representações) apresenta-se como é que as tabelas de frequências foram introduzidas e trabalhadas nas aulas e mostra-se como foi desenvolvida a função cumulativa, em particular, as relações estabelecidas desta função com outras representações. Na segunda secção (reflexão sobre o tópico na aula) inclui-se reflexões da professora sobre o ensino e aprendizagem do tópico. Na terceira secção (discussão do tópico) realiza-se uma discussão em torno da *Organização e interpretação de dados*, guiada pelas quatro dimensões do modelo do conhecimento didático do professor em Estatística.

6.3.2.1 Representações

A tabela de frequências foi uma das representações mais usadas como síntese dos dados. Os primeiros dados, incluídos em tarefas, a serem organizados em tabelas de frequências foram qualitativos e todos os restantes foram de natureza quantitativa. Lia propôs várias tarefas que envolviam uma representação gráfica (pictograma, diagrama circular, gráfico de barras, histograma, diagrama de caule e folhas) cujo objetivo principal era chegar à representação tabular correspondente, ou seja, os dados teriam de ser representados numa tabela com frequências (absolutas e relativas) e, em alguns casos, as tarefas pediam também as frequências acumuladas.

Tabelas de frequências

As primeiras duas tarefas referidas nesta secção incluem dados representados num gráfico de barras e a terceira inclui dois conjuntos de dados quantitativos contínuos.

Tarefa *Frequência com que os alunos de uma escola comem fast food* (t5)

Uns minutos depois de ter distribuído esta tarefa aos alunos, Lia pergunta-lhes se alguma vez tinham ouvido falar de frequências absolutas acumuladas, ao que eles disseram que não. Pede ao aluno André para ler em voz alta a definição que ela projetou num *powerpoint*: “Frequência absoluta acumulada de um valor da variável é a soma das frequências absolutas dos valores da variável desde o primeiro até ao valor pretendido, inclusive”. Depois, há a seguinte interação:

> **Professora:** Alguém consegue, a partir daqui, perceber?
> **Alunos:** Não (muitos alunos indicam não perceber)
> **Professora:** Vamos por partes… Façam o seguinte… Comecem por fazer uma tabela com as frequências absolutas simples… números de vezes por semana que os alunos comem *fast food*… Olhando para o gráfico, há 0, 1, 2, 3, 4, 5 e 6 vezes. A frequência simples é o *ni*. A frequência acumulada *Ni*. Toda a gente conseguiu preencher a primeira coluna?... [faz a tabela no quadro]… Toda a gente conseguiu?… Quantos alunos comem *fast food* zero vezes por semana?
> **Vários alunos:** 25
> **Professora:** E a seguir?
> **Alunos:** 75, 50, 200…
> **Professora:** [Observando a tabela] Vamos fazer a leitura da frequência acumulada… Quem ainda não percebeu o que é a frequência acumulada coloque o dedo no ar… Ah! Então, vou-vos explicar… Para uma vez [na coluna do *Ni*]… quantos alunos comem por semana 0 ou 1 vez?… Tenho de acumular com o que está para trás!
> **Alguns alunos:**100.
> **Professora:** Fica-se com 100 [25+75]. 100 alunos comem 0 ou 1 vez… incluindo uma, *fast food* por semana. O que representa o 150 [*Ni*], na coluna das acumuladas? É o número de vezes que os alunos comem *fast food* 2, 1 e 0 vezes, ou seja, menos de 2 vezes, incluindo o duas vezes…
> **Rafael:** Stora o que significa o último 500 na coluna das frequências acumuladas?
> **Professora:** Então o que vai dar o total?... Não faz sentido somarmos os valores da coluna das frequências acumuladas porque já o estamos a fazer gradualmente… Dúvidas?... Para que serve a coluna das acumuladas, então?... (Ninguém responde, a maioria está ainda a preencher a tabela).
> A coluna das frequências acumuladas tem as somas feitas… Se não houvesse essa coluna… teriam de as fazer [somas] dependendo da pergunta que vos fizessem… (…) O que podem dizer [com a informação] da tabela? Os alunos comem muitas vezes *fast food* ou poucas? O que acham?
> **Luís:** Há muitos alunos que comem 3 vezes por semana.
> **Luzia:** A moda é 200!
> **Professora:** Ah! E o que é a moda?
> **Luzia:** O valor que aparece mais vezes…
> **Professora:** Iremos relembrar a mediana e a média nas próximas aulas. Quanto à moda, parece estar na moda comer três vezes por semana! (observação 2)

Lia usou esta tarefa para introduzir as frequências acumuladas. Como vários alunos indicaram que não tinham percebido a definição que ela pediu que fosse lida no *powerpoint* que produziu, Lia optou por ir preenchendo a tabela em conjunto com os alunos, esclarecendo símbolos e algumas quantidades usadas. Quando assinalou um valor da coluna das frequências acumuladas associou-o a uma soma parcial, descrevendo-o como um valor que corresponde a um acontecimento que inclui os respetivos valores da variável estatística. Indicou a utilidade das somas parciais da coluna das frequências acumuladas para se responder a questões de forma expedita. Ao pedir uma apreciação global à informação da tabela sobre a situação alimentar dos alunos, há uma aluna que responde, identificando imediatamente a moda da distribuição.

Na aula, Lia voltou a inquirir os alunos para se certificar se a tabela preenchida tinha sido compreendida por todos, enquanto mostrava, num *powerpoint*, a tabela de frequências absolutas e relativas (simples e acumuladas – figura 50):

N.º de vezes por semana	Frequência absoluta		Frequência relativa	
	simples (n_i)	acumulada (N_i)	simples (f_i)	acumulada (F_i)
0 vezes	25	25	0,05	0,05
1 vez	75	100	0,15	0,2 (= 0,05 + 0,15)
2 vezes	50	150	0,1	0,3 (= 0,2 + 0,1)
3 vezes	200	350	0,4	0,7 (= 0,3 + 0,4)
4 vezes	95	445	0,19	0,89 (= 0,7 + 0,19)
5 vezes	50	495	0,1	0,99 (= 0,89 + 0,1)
6 vezes	5	500	0,01	1 (= 0,99 + 0,01)
Total (N)	500		1	

Figura 50: Tabela de frequências (absolutas e relativas).

Professora: Na parte das frequências absolutas simples e acumuladas. Por exemplo, o que é que representa o 150?... Rui?
Rui: Os que comem duas vezes ou menos.
Professora: Ou seja, o número de vezes que comem por semana é dois, uma ou zero. Como chegaste ao 150 [da coluna das frequências absolutas acumuladas]?
Rui: Somei 25 mais 75 mais 50.
Professora: Ou seja, o 150 resulta… acumulaste 50 com 25 mais 75 [da parcela anterior]. Luzia, o que significa os 445?
Luzia: O número de alunos que comem 4 vezes ou mais.
Professora: Se fosse 4 ou mais, terias de fazer a contagem com tudo para a frente! Querias dizer 4 ou menos… Podemos dizer antes que são os que comem no máximo 4, ou 0,1,2,3 e 4 vezes. Também há as frequências relativas simples e acumuladas… portanto frequências em percentagem… Por exemplo, 50% onde estão? São os de 2 ou menos vezes, certo?!... Então, vamos lá!... Marisa, olhando para a tabela, qual a percentagem de alunos que comem… no máximo três vezes?

Marisa: 70%.
Professora: Somaram os valores, não é? É para isso que serve a acumulada… (observação 2)

Nesta intervenção, Lia acaba por ajudar os alunos a simplificar o modo como descrevem os valores das frequências acumuladas (absolutas ou relativas), ou seja, cada soma parcial, no fundo, corresponde a um acontecimento que inclui, no máximo, os respetivos valores da variável estatística. Também corrigiu o raciocínio da aluna Luzia, que revelou não estar a perceber totalmente como é que os valores das frequências acumuladas eram formados.

Tarefa *Evolução da rede de saneamento* (t12)

Depois de ter selecionado esta tarefa do manual adotado, Lia pediu aos alunos para que a lessem com muito cuidado e respondessem às três questões do enunciado. As duas primeiras implicavam recolher informação junto do gráfico fornecido e a terceira a construção de uma tabela de frequências absolutas e relativas. Vários alunos chamaram a professora aos seus lugares porque não sabiam como fazer para responder às duas primeiras alíneas. Eles pareciam também não saber expressar o que pretendiam através do gráfico. Ao aperceber-se que a maioria da turma não estava a conseguir avançar, há o seguinte intercâmbio:

Professora: Em 2002. Qual foi a extensão?
Lucas: 54.
Professora: E em 2003?
Rafael: Foi 20.
Professora: Ah! Foi 74 menos quê?… Porquê? O que se está a passar?
Rafael: Está a acumular!
Professora: Qual é o gráfico de barras que está aí representado? É o gráfico de frequências absolutas…
Rafael: Acumuladas!
Professora:… Acumuladas! O gráfico fala de uma evolução… certo?! Para sabermos em que ano houve uma maior evolução da rede de saneamento temos de subtrair…
Rafael: Subtrair os anos…
Professora: Vamos subtrair o 74 a 54… Que mais?... O 112 a?
Rafael: A 186. É?
Professora: O 112 a 74.
Rafael: Ah!… No ano 2002 há 54 e no ano 2003 há 74, que é 54 mais 20… E o 20 é 74 menos 54… É complicado explicar…
Telma: Posso explicar?… Por exemplo… há uma evolução [acréscimo] de 20 em 2003…
Rafael: Em cada ano tenho de subtrair o anterior!

> **Professora:** No total há quantos quilómetros de extensão? 315 quilómetros… tal como a última barra do gráfico indica! (observação 2)

Nesta interação, foi feita uma ligação entre o diagrama de barras do enunciado e as frequências acumuladas, acabando por ser explicado o método para se saber exatamente a evolução (em metros) da rede de saneamento nos anos em que ela decorre com a participação dos alunos Rafael e Telma.

Após esta intervenção, há alguns alunos que optam por começar a resolução da tarefa pela terceira alínea, ou seja, pela construção da tabela para só depois responderem às restantes questões. Lia ainda detetou algumas dificuldades no preenchimento da tabela, que pedia as frequências absolutas e relativas. A dificuldade dos alunos no preenchimento da tabela prendeu-se fundamentalmente com a passagem das frequências absolutas acumuladas para as frequências absolutas simples. Lia obteve as respostas corretas nas restantes alíneas, assim que questionou alguns deles, no momento da correção da tarefa em grande grupo.

Tarefa *Resultados do campeonato de atletismo nos 80 metros* (t7)

Esta tarefa tinha a particularidade de fornecer dados individuais contínuos. Tinha também a especificidade de pedir que se fizesse o polígono de frequências absolutas e o de frequências absolutas acumuladas, que nunca tinha surgido na aula. Para fazer a tabela pedida no enunciado, os alunos teriam de agrupar os dados em classes.

Para que não se perdesse muito tempo nesta tarefa, Lia preparou um *powerpoint* em que mostrou como aplicar a regra (que vinha descrita no enunciado) que permitia descobrir quantas classes formar (figura 51).

> **Quantas classes se devem definir?**
>
> Existe uma fórmula que nos fornece uma indicação do número de classes a considerar:
>
> O número de classes, k, a considerar numa amostra de dimensão n é o menor número inteiro tal que
>
> $$2^k \geqslant n$$

Figura 51: Excerto retirado da aula.

Além disso, mostrou num *slide* quais seriam as classes para a situação descrita na tarefa. Os alunos copiaram essas classes e foram completando, nos seus lugares, uma tabela de frequências que incluía frequências absolutas e relativas e as respetivas acumuladas. Assim que a maioria indicou ter completado a tabela, Lia mostrou um *slide* com a mesma totalmente preenchida, questionando apenas os alunos sobre se tinham obtido as mesmas contagens e resultados. Nesta interação, não se registou dificuldades nos alunos.

Também, através do mesmo *powerpoint*, a professora apresentou as restantes respostas às questões da tarefa, ou seja, um polígono de frequências desenhado sobre um histograma de frequências absolutas e outro sobre um histograma de frequências acumuladas, tendo o cuidado de dedicar algum tempo a explicar quais os pontos que deveriam ser ligados em cada circunstância (figura 52). Nestas situações, não foi feito nenhum comentário à distribuição de dados nem ao facto de se ter optado por juntar os dados numa única distribuição, uma vez que, na tarefa original, os dados aparecem distinguidos pelo sexo dos atletas.

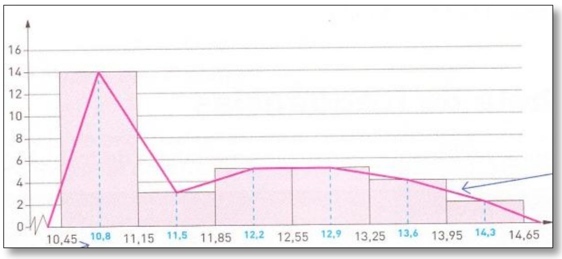


Figura 52: Polígonos de frequências absolutas e acumuladas, respetivamente.

Função cumulativa

A tarefa *Despesas em transporte de 20 trabalhadores* (t6) foi a única proposta na aula abordando, de modo explícito, a função cumulativa. Para introduzir a função cumulativa, Lia usou os dados (quantitativos discretos) de uma tarefa que já tinha sido trabalhada na aula (t5), tal como se mostra de seguida.

A professora aproveitou os dados discretos da tarefa *Frequência com que os alunos comem fast food* (t5) para exemplificar "como fazer a passagem de um gráfico de barras de frequências acumuladas para o gráfico da sua função cumulativa" (observação 2), com indicação da expressão analítica desta função e sua interpretação. Na aula, Lia fez sobretudo uma leitura concisa ao *slide* que preparou com essas representações e descrições (figura 53).

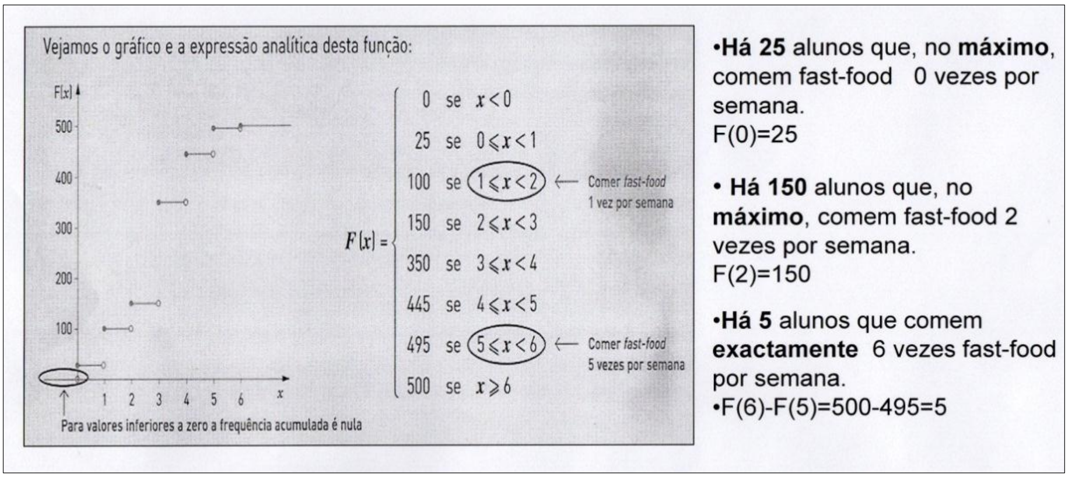


Figura 53: Excerto com duas representações de uma função cumulativa.

Tarefa *Despesas em transportes de 20 trabalhadores* (t6)

Nesta tarefa, a professora pediu aos alunos para se guiarem pelas representações que tinha mostrado com os dados da tarefa *Frequência com que os alunos de uma escola comem fast food* (t5), que tinha sido trabalhada anteriormente, para a determinação da expressão analítica da função cumulativa da atual tarefa (t6). Mais especificamente, pretendia que os alunos escrevessem a função cumulativa *F(x)* analiticamente a partir do seu gráfico. Depois de circular pelos lugares, detetou algumas dificuldades nalguns alunos em expressar quais os ramos e os respetivos valores acumulados, optando por exibir, na tela, a função cumulativa pedida na tarefa proposta, que tinha preparado para o efeito num *slide* (figura 54):

$$F(x)=\begin{cases}0 & \text{se } x<14\\ 4 & \text{se } 14\leqslant x<15\\ 4+7 & \text{se } 15\leqslant x<16\\ 4+7+5 & \text{se } 16\leqslant x<17\\ 4+7+5+2 & \text{se } 17\leqslant x<18\\ 4+7+5+2+1 & \text{se } 18\leqslant x<19\\ 4+7+5+2+1+1 & \text{se } x\geqslant 19\end{cases} \quad \text{ou} \quad F(x)=\begin{cases}0 & \text{se } x<14\\ 4 & \text{se } 14\leqslant x<15\\ 11 & \text{se } 15\leqslant x<16\\ 16 & \text{se } 16\leqslant x<17\\ 18 & \text{se } 17\leqslant x<18\\ 19 & \text{se } 18\leqslant x<19\\ 20 & \text{se } x\geqslant 19\end{cases}$$

Figura 54: Função cumulativa.

Professora: Para vos ajudar, indico as somas… na expressão analítica. Agora já podem dizer: quantos são os trabalhadores que gastam no máximo 18 euros?
Martim: 19.
Professora: É o F(19). Ivo, qual é a questão seguinte?
Ivo: Quantos trabalhadores gastam por semana exatamente 16 euros?
Professora: Como é que fizeste Ivo?
Ivo: Acho que é F(16) menos F(15)…
Professora: Exato, 16 menos 11, que dá 5 trabalhadores! (observação 2)

Apesar de ter ainda despendido algum tempo a observar como é que os alunos se estavam a sair na determinação da expressão analítica da função cumulativa e a explicar individualmente, a alguns, como é que se fazia, optou por apresentá-la, através do *slide* que trazia preparado, em duas formas conjuntamente: com as somas parciais à vista, por ramos, e depois com o resultado final desses valores, também, por ramos. Os alunos intervenientes na interação anterior revelaram saber usar a função para responder acertadamente às questões.

6.3.2.2 Reflexões sobre o tópico na aula

Lia referiu que este tópico dizia essencialmente respeito à leitura de informação numa representação gráfica e à construção de tabelas de frequências. Acrescentou que, “na organização de dados, não se fala na interpretação, nem na análise” (entrevista 2). Assim sendo, Lia esperava que este tópico fosse acessível e considerava que os alunos, de uma maneira geral, não iriam revelar muitas dificuldades em trabalhá-lo. Lia realçou, ao longo das três entrevistas realizadas neste estudo, que o enfoque, neste tópico, estava na sumarização dos dados nalguma representação tabular ou, equivalentemente, num diagrama de barras ou histograma. Ao relembrar a tarefa que foi resolvida na aula que se relacionava com alcançar a função cumulativa (t6), considerou que, através destas situações, os alunos reconheceram "a função cumulativa mais como uma alternativa de representar o diagrama de barras de frequências acumuladas" (entrevista 2). Também declarou a sua opinião que a função cumulativa na Estatística do ensino secundário não tinha "aplicações" (entrevista 2), considerando que os próprios manuais expressavam essa perspetiva por colocarem poucas tarefas sobre este assunto. Assim, neste tópico, a maioria das tarefas que recomendou tinha o propósito principal de fazer com que os alunos consolidassem a construção de tabelas.

Lia selecionou a tarefa *Evolução da rede de saneamento* (t12) do manual escolar para propor na aula por esta incorporar um gráfico de barras (com as barras organizadas em sentido ascendente) que fazia alusão às frequências acumuladas e, por conseguinte, poder estar associada à função cumulativa, o último conceito estudado na aula. Lia relembrou a atuação dos alunos da seguinte forma:

> [Foi proposta] Logo após se ter dado a tabela de frequências acumuladas e a função cumulativa… Vários alunos não perceberam que era gráfico… cumulativo… No início tiveram dificuldades, porque achavam que isto era relativamente… porque… tinham de ver de um ano para o outro, o quanto é que estava a crescer [a rede de saneamento]. No início estavam a interpretar que tinha crescido 74 quilómetros, mas não! [valor que coincidia com a altura da segunda barra]… tinham de ver as diferenças (…) Nesta parte [organização de dados], era tudo muito direto… com construção de tabelas… Lá está! Não tinha nada de interpretar… não havia interpretações, nem explorações… que é onde eles [alunos] têm maior dificuldade. (entrevista 2)

Neste excerto, Lia recordou as dificuldades que a maioria dos alunos estava a sentir na interpretação da representação inserida na tarefa. Os alunos não estavam a associar o gráfico com o que tinham acabado de aprender sobre frequências acumuladas e função cumulativa. Simultaneamente, Lia deixou transparecer a sua visão geral sobre o tópico, tratando-se de um tópico com uma finalidade própria, que tem a ver com os alunos desenvolverem, sobretudo, a capacidade de construir representações tabulares como síntese dos dados fornecidos, não havendo necessidade de recorrerem a explorações ou interpretações.

Ao relembrar a atuação na aula do aluno Rafael que, na altura, entendeu rapidamente que a representação gráfica da tarefa t12 dizia respeito a frequências acumuladas, Lia mostrou a sua surpresa: "Fiquei surpreendida com o Rafael naquelas questões… Ele percebeu logo o gráfico, que correspondia a frequências acumuladas" (entrevista 2), confessando, adicionalmente, que o Rafael nunca participou tanto nas aulas de Matemática como o fez na Estatística. Considera que o Rafael, e mesmo o Mário, tiveram uma prestação bastante positiva na Estatística contrariamente ao que mostraram nos restantes temas matemáticos, acrescentando: "Há alunos que não têm grande jeito para a Matemática, mas gostam muito de desbobinar as suas opiniões… [na Estatística] o Rafael passou a ir ao quadro e a falar sobre os assuntos" (entrevista 3).

Lia parece deter a ideia que alunos com dificuldades em Matemática, e também com um certo à vontade em comunicar as suas opiniões sobre assuntos diversos, têm, no tema da Estatística, uma oportunidade de mostrar essas suas capacidades, por exemplo, na interpretação das situações trabalhadas.

6.3.2.3 Discussão do tópico

CETE. Lia deu particular destaque à construção de tabelas de frequências a partir de representações gráficas diversificadas, dado terem sido resolvidas várias tarefas com esse propósito. O *menu* da Estatística da calculadora gráfica não foi usado neste tópico. Teve algumas interações com os alunos para que eles ficassem a perceber como preencher uma tabela de frequências absolutas acumuladas e a sua utilidade para responder a algumas questões sobre os dados. O mesmo sucedeu quando abordou o gráfico de barras de frequências acumuladas para associá-lo à função cumulativa. Houve alguns momentos expositivos, nomeadamente, quando a professora mostrou como se aplicava a regra para se determinar o número de classes a definir para um conjunto de dados quantitativos contínuos, quando referiu como é que os polígonos de frequências eram formados, e quando introduziu a função cumulativa. Nesses momentos Lia teve uma menor interação com os alunos. Foi curioso observar que, num momento específico em que a professora provocou os alunos, questionando-os sobre reações relativamente à tabela construída, nomeadamente, o que é que ela transmitia sobre a situação em causa, a aluna que respondeu apontou que visualizava a moda da distribuição na tabela. Apesar de a professora não ter dado seguimento à situação, esta parece espelhar um pouco o potencial de questões que envolvam dados e o seu contexto para orientar os alunos para análises mais ricas. De uma maneira geral, não se registou uma exploração aprofundada das relações entre as diferentes representações que surgiram na aula, intuindo-se sobre o que estas transmitiriam sobre os dados em estudo.

CA. De acordo com a experiência de Lia, este tópico é acessível aos alunos. No entanto, a professora foi detetando algumas dificuldades na leitura das tabelas de frequências. Encontrou também dificuldades na leitura de um gráfico de barras (frequências acumuladas) de modo a responder às questões colocadas na tarefa. Neste tópico, não houve muitas situações em que os alunos tivessem de explicar as suas respostas ou de fazer comentários; por conseguinte, o acesso ao raciocínio dos alunos foi limitado. Lia considera que há alunos mais fracos em Matemática que acabam por ter um desempenho razoável em Estatística devido a poderem mostrar mais as suas capacidades comunicativas e opiniões sobre os assuntos tratados.

CE. Lia conhece várias representações e associa função cumulativa a frequências acumuladas. O modo como as tarefas foram exploradas nas aulas não permitiu capturar outros aspetos relevantes do conhecimento de Estatística da professora, no ensino deste tópico.

CC. Neste tópico, Lia tinha claramente o objetivo de trabalhar a construção da representação tabular (frequências absolutas, relativas e também acumuladas) e apresentar essa representação como forma de sintetizar os dados. Na sua prática admitiu leituras pontuais de representações tabulares, mas sem ter em vista análises estatísticas dos dados. A função cumulativa surgiu também relacionada com frequências acumuladas de dados quantitativos discretos.

6.3.3 Medidas de localização

It is impossible to consider center without also considering spread,
as both ideas are needed to find meaning in analyzing data
Garfield e Ben-Zvi (2008, p.188)

Este tópico é analisado em três secções. Na primeira secção (média, moda e quartis) começa-se por fazer referência às situações em que as medidas de localização surgem na aula e é também analisado como é que essas medidas foram desenvolvidas na prática da Lia. Na segunda secção (reflexões sobre o tópico na aula) apresenta-se algumas reflexões da professora sobre o ensino e aprendizagem do tópico. Na terceira (discussão do tópico), e última secção, mostra-se uma discussão em torno da *Organização e interpretação de dados*, guiada pelas quatro dimensões do modelo do conhecimento didático do professor em Estatística.

6.3.3.1 Média, moda e quartis

As medidas de localização foram trabalhadas na aula com tarefas que providenciavam conjuntos de dados numéricos, alguns deles agregados em classes. Na maioria das situações e tarefas com dados agrupados apresentadas, as classes surgiam já definidas. Apenas na tarefa *Resultados do campeonato de atletismo nos 80 metros* (t7) foi solicitada a determinação das classes através da fórmula que vinha incluída na tarefa. A visualização das medidas estatísticas moda, mediana e quartis em tabelas, diagramas de barras ou histogramas surgiu em tarefas que envolviam dados agrupados em classes. As medidas de localização foram também abordadas em tarefas que envolviam a determinação das medidas de dispersão. Lia não deu uso à calculadora gráfica nesta temática; contudo, quando abordou as medidas de dispersão, informou que a instrução que permitia obter o valor do desvio-padrão servia também para encontrar os valores, mínimo, máximo, média, mediana e os restantes quartis no ecrã da calculadora. Indicou ainda que a moda surgia nessa lista de estatísticas sumárias na calculadora gráfica da *Casio*, mas não na da *Texas Instruments.*

Média

Nesta subsecção, mostra-se como é que a noção de média surgiu e foi desenvolvida na aula, sobretudo, no tratamento de conjuntos de dados quantitativos discretos (nas tarefas t17 e t23). Também se expõe a média perante conjuntos de dados quantitativos contínuos através da situação *Zapping da Adriana* e da tarefa *Inquérito a 500 alunos sobre televisão* (t20).

Tarefa *Qual a melhor turma, A ou B?* (t17)

Esta tarefa foi proposta pela professora para introdução das medidas de localização. Lia, ao ler o enunciado na aula, começou por fazer um pequeno esclarecimento relativamente à leitura das unidades dos dados, dizendo "por exemplo, 190 pontos significam 19 valores" e pediu aos alunos para se debruçarem na primeira alínea, no cálculo da média, moda e mediana. No momento de discussão e correção da tarefa, há a seguinte interação:

> **Professora:** As médias são muito diferentes?
> **Alguns alunos:** Não.
> **Professora:** Portanto vocês têm as classificações, calcularam a média, que é uma medida de localização ou de tendência central… As medidas de localização são medidas que nos ajudam a tirar conclusões acerca dos resultados. Não tiveram dificuldades em calcular a média, pois não?
> **Vários alunos:** Não.
> **Professora:** Como é que fizeram?
> **Professora e alunos** (dizem em simultâneo): Somamos os dados e dividimos pelo número total.
> **Professora:** A média representa-se por *x* barra (com a barra em cima do *x*)… (lê em voz alta, através do *powerpoint* exibido) 'é uma das medidas mais usadas para se resumir e comparar distribuições, por um lado… Mas, por outro lado,… é também afetada pela existência de alguns valores extremos'. O que é que isso quererá dizer?… Quem me sabe dizer?... Quando digo que ela pode ser afetada pela existência de valores extremos, o que é que isto quer dizer?... Diz lá alto, Martim.
> **Martim:** Quando há valores muito grandes ou muito pequenos (referindo-se a dados com valores muito altos ou muito baixos), a média pode dar um valor que não é perto da realidade, é mais ou menos…
> **Professora:** (Abana a cabeça para indicar a sua concordância com a resposta do Martim e diz) Quando há dados muito altos ou muito baixos, a média pode realmente não ser representativa do conjunto de dados, certo? Vamos agora comparar as médias das duas turmas. (observação 3)

A média foi identificada como uma medida de localização, ou de tendência central, e foi calculada pelo seu conhecido algoritmo. Lia também abordou a pouca resistência da média face a valores extremos. Nesta interação, os valores extremos foram associados a dados com valores muito baixos ou muito altos. Ao concordar com a explicação do Martim quanto à influência dos valores extremos no cálculo da média, Lia também interligou a ideia da média como uma medida sensível, face à presença de algum valor extremo, com um valor de média pouco "representativo do conjunto de dados" (observação 3).

De seguida, mostra-se como é que Lia continua a interação com os alunos sobre os resultados da média obtida para cada turma. Pelo facto de estes valores serem próximos para as duas turmas, Lia transmite-lhes a ideia que não se devem precipitar a retirar conclusões na comparação destes valores. E também parece considerar que a variabilidade presente nos dados de cada turma não é refletida nos valores de cada média. Assim, apela à necessidade de se observar o comportamento dos dados e de se calcular medidas estatísticas adicionais:

> **Professora:** Na turma A [a média] deu 78 pontos, foi isso? E na turma B?
> **Vários alunos:** 76,3 pontos.
> **Professora:** A diferença destas médias não é muita, pois não?
> **Alguns alunos:** Não é muito.
> **Professora:** Será que podemos dizer que as duas turmas tiveram resultados idênticos?
> **Mário:** Mas a turma B tem mais um aluno!
> **Professora:** Sim, independentemente disso, não afeta o cálculo… Imaginem que só sabiam os resultados das médias [desconhecendo os dados]. Isso quer dizer que as duas turmas têm o mesmo tipo de notas?
> **Alguns alunos:** Não.
> **Professora:** O que posso concluir?... Que não há uma melhor que a outra?
> **Alguns alunos:** Não.
> **Professora:** Então a média não chega para se tirar uma conclusão! Temos de calcular mais algumas medidas estatísticas que nos ajudem a tirar conclusões. Que posso concluir sobre a média? …
> **Marisa:** A média não é certa…
> **Professora:** Certo? O que queres dizer?
> **Marisa:** Que nem sempre a média serve para tirar conclusões…
> **Professora:** Nem sempre serve… Poderá não chegar para tirar conclusões, depende dos dados que tenho, correto? (observação 3)

A intervenção do aluno Mário parece indicar que não se pode comparar os valores das médias pelo facto de uma turma ter mais um aluno do que a outra. A professora reage dizendo que a sua observação não interfere no cálculo da média. Lia também passa a ideia que os alunos não se devem basear apenas na média para argumentarem as suas escolhas, nomeadamente em situações em que as médias são próximas. E nesta tarefa concreta sugere, adicionalmente, que se observe a variabilidade dos dados; contudo, a descrição dessa variabilidade não surge explicitamente durante a exploração e resolução da tarefa.

Situação *Zapping da Adriana* (anexo 4) e Tarefa *Inquérito a 500 alunos sobre televisão* (t20)

Antes de mostrar a situação *Zapping da Adriana*, que incluía dados agrupados em classes, Lia justificou a necessidade de se recorrer ao agrupamento de dados quando há "um número muito grande de dados". Também advertiu os alunos sobre a impossibilidade de se obter um "valor exato de média para dados agrupados em classes, só se tivermos os dados todos…" (observação 4). Ou seja, informou que, com os dados agregados, poderiam obter somente um valor aproximado da média e que, para este cálculo, os alunos precisariam de determinar a marca de cada classe:

> **Professora:** O que é a marca de cada classe? Por exemplo, na classe [0, 10[… qual é a marca desta classe?
> **Eva:** É 10.
> **Professora:** Isso é a amplitude… A marca é o valor médio da classe…
> **Martim:** Ah…
> **Eva:** Então é 5. (observação 4)

Com a situação *Zapping da Adriana* (figura 55), Lia apresentou o cálculo da média através do algoritmo usual, que dizia respeito à divisão entre soma dos produtos das marcas de cada classe pela sua frequência e número total de dados. Durante a interação entre a professora e alunos, alguns alunos manifestaram não perceber o significado da coluna que incorporava as classes da tabela, tendo em conta o contexto da situação. Para esclarecer os alunos, Lia fez, por várias vezes, leituras da tabela. Como exemplo, apresenta-se o que ela disse quanto à primeira classe: "A Adriana em 9 dias fez *zapping* entre 0 e 5 vezes… 2,5 [que é a marca desta classe] é um valor que se arranjou para representar a classe e para se poder calcular a média" (observação 4). Neste exemplo, o cálculo do valor aproximado da média acabou por ser corretamente alcançado pela maioria dos alunos.

Classes	Frequência absoluta – f_i	Valor central da classe – x_i	$f_i\,x_i$
[0, 5[	9	2,5	22,5
[5, 10[	7	7,5	52,5
[10, 15[	4	12,5	50
[15, 20[	6	17,5	105
[20, 25[	5	22,5	112,5
Total	$\Sigma f_i = 31$		$\Sigma f_i\,x_i = 342{,}5$

$$\overline{x} \approx \frac{342{,}5}{31}$$

$$\overline{x} \approx 11{,}05$$

Figura 55: Tabela com os dados do *Zapping* da Adriana e cálculo do valor da média.

Num outro momento, enquanto os alunos trabalhavam na tarefa *Inquérito a 500 alunos sobre televisão* (t20), Lia ia circulando pela sala de aula e foi-se apercebendo que havia alunos que estavam a ter dificuldades em determinar a média pedida. Mesmo os que fizeram o cálculo, nos seus lugares, desejavam, ainda assim, confirmar junto da professora as suas respostas para terem a certeza de que estavam a fazer bem. Por exemplo, o aluno Carlos mostrou que tinha feito o cálculo da média justamente com a informação do gráfico circular (somou as percentagens de cada uma das seis áreas do gráfico circular e dividiu esse valor por 6):

Professora: (ao tentar perceber o cálculo do aluno diz) Fizeste a classe média?! [ri-se]
Carlos: O cálculo deu 0,16… salvo erro… Este valor está na classe [15, 30[…
Professora: Já reparaste no cálculo que fizeste?... O que significam os 8% (que corresponde à classe 0-15 no diagrama circular)?... 8% de 500… Quantos são?... Já calculaste?… [para a tabela]… São 40 alunos (Lia procura essa resposta na tabela do manual do Carlos que já tinha sido preenchida por ele)… Ali… estavas a calcular a média dos dados em percentagem… do diagrama circular?... A pergunta é… a média dos tempos gastos… é a variável em estudo… média dos tempos gastos diariamente a ver televisão pelos alunos inquiridos… certo? … No total, são 500 [alunos]… Foste calcular a média das percentagens, não foi?
Carlos: Sim.
Professora: Entendes agora o que fizeste?… Não é o pedido! Entendes? (observação 5)

Lia fez também uma leitura da tabela para o Carlos entender que informação deveria escolher da tabela para fazer o algoritmo da média, nomeadamente, a marca e as frequências absolutas. Vários alunos foram ao quadro apresentar a resolução das diferentes alíneas da tarefa. No que diz respeito à média, foi apresentado o seu algoritmo e resultado (figura 56). Não foram tecidos comentários a este resultado.

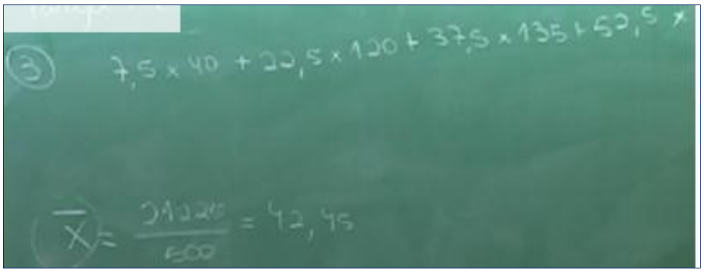


Figura 56: Cálculo do valor da média (t20).

Lia detetou que o Carlos estava a ter dificuldades em perceber que dados usar para calcular a média pedida, quando tinha à sua frente um gráfico circular e uma tabela que já tinha sido corretamente preenchida por ele no lugar. Essa tabela permitia-lhe observar os dados registados no gráfico circular agora em frequências absolutas, relativas e relativas acumuladas. O aluno achava que, para determinar a média, bastava somar as percentagens de cada setor do gráfico circular, dividir pelo número total de setores, e, ainda, dividir esse resultado por cem. Lia ajudou-o a perceber que o que se pretendia fazer era uma média dos tempos gastos a ver televisão e que, para isso, era preciso ter em conta os diferentes tempos observados (intervalos) com as respetivas frequências absolutas (que o aluno tinha determinado e registado corretamente na tabela do enunciado). O aluno não teve dificuldades em preencher a tabela mas apresentou como média, uma média feita com as frequências relativas. Assim, notoriamente, revelou dificuldades em interpretar os dados e em usar as marcas das classes de modo a aplicar o algoritmo corretamente.

Tarefa *Qual das duas gémeas a melhor aluna?* (t23)

Os alunos determinaram os valores da média, moda e mediana para cada uma das irmãs, com o apoio da calculadora gráfica, mas sem usarem o menu *STAT*. De uma maneira geral, todos chegaram aos mesmos resultados, ou seja, a valores de média, moda e mediana de treze (valores) para cada uma das irmãs.

Durante o desenvolvimento da tarefa, a uma dada altura, Lia pretendia que os alunos chegassem à conclusão que as medidas de localização calculadas, nomeadamente, a média, que dera igual para as duas irmãs, era uma medida que não descrevia a variação que existe nas notas de cada irmã. Como consequência, apontou a necessidade de se recorrer a outras medidas para se dar resposta à questão do enunciado: "Qual das duas gémeas é a melhor aluna?":

> **Professora:** (…) Apesar de a média ser igual, o desempenho [das duas alunas] foi diferente… Imaginemos então que não sabiam nada sobre as notas delas. Imaginemos que não vos tinha dado as notas delas. Só [sabiam] que a média, moda e mediana de cada uma delas era treze ou sabiam que a média das duas era igual 13. Será que vocês diriam logo assim: 'Se a média é igual nas duas… as duas merecem a mesma nota!'. É isso?
> **Alguns alunos:** Não.
> **Professora:** Neste caso, vocês verificam que não merecem… Isto quer dizer o quê? A média nem sempre… Ao calcular a média, o que podemos concluir?
> **Mário:** Os resultados [notas] obtidos podem não corresponder à média…
> **Professora:** Isto quer dizer então… Quando for a dar as vossas notas finais… eu não posso só olhar para a média, pois não? O que acham?
> **Sofia:** Só para a média… não!?
> **Professora:** Tenho de olhar para o conjunto de dados… ver se existem notas muito baixas ou notas muito altas… para ver se existe muita dispersão dos dados… É ou não é? Neste caso… a Joana, se calhar, merecia mais, apesar da média dela ser treze, do que a outra irmã. Ela [Mariana] teve um 19, um 7 e um 8 e … um 18 e depois teve as notas medianas à volta do 13… Ao analisarmos as notas, estamos a ir mais além… Joana relativamente à Mariana é uma aluna mais… mais o quê?
> **Eva:** Consistente…
> **Professora:** (Acenando a cabeça afirmativamente diz) Mais constante… relativamente às matérias que são dadas… Independente das dificuldades vai-se mantendo em torno do treze… que é a média… E a Mariana, vai-se mantendo à volta do treze?
> **Alguns alunos:** Não.
> **Professora:** [Por exemplo a] Mariana tem melhores notas que treze… São muito desviadas do treze, da média. Agora estou a falar das notas relativamente à média. Porque é que ela teve um 7 e depois um 17?... Os dados permitem questionar isso (…) Vamos continuar… Então verificámos [que], se calhar, precisamos de calcular mais algumas medidas para estudar os dados… Vamos recorrer às chamadas medidas de dispersão… que permitem estudar a variabilidade dos dados… (observação 6)

Com esta interação, a professora parece ter o propósito de levar os alunos a pensar que, para tomarem decisões quanto à situação em estudo, não se devem focar apenas nos resultados das medidas calculadas (nomeadamente da média), mas também terem o cuidado de entender essas medidas tendo em conta a distribuição que têm em mãos. A professora também pretendia que os alunos soubessem criticar os resultados das medidas obtidas, fazendo o exercício de duas formas: com o conhecimento dos valores dos dados ou ignorando-os. Quando o aluno Mário se refere à não correspondência entre as notas e a média obtida, a professora aponta a necessidade de se olhar para certos aspetos particulares nos dados, mas não corrige a resposta inadequada do aluno. Os aspetos apontados por Lia diziam respeito à averiguação de quanto os dados fornecidos (resultados dos testes de cada gémea na disciplina de Matemática) estavam ou não próximos da sua média, situação que foi realizada para cada gémea, embora sem se especificar os valores numéricos envolvidos nesses desvios. Foi observada a necessidade de Lia combinar a medida de centro média com a variabilidade presente dos dados na análise que fez, e de ponderar a possibilidade de se retirar conclusões com essa medida de localização.

Moda

Na tarefa *Qual a melhor turma, A ou B?* (t17), a moda é apresentada pela professora como "o elemento do [conjunto de dados] com maior frequência" (observação 3). Na interação que houve entre professora e alunos, quando é corrigida a alínea que pede a identificação da moda, a professora destaca a inconveniência de se retirar conclusões sobre a melhor turma através desta medida:

> **Professora:** O que é a moda? Como se representa? Por *Mo*, é o elemento da amostra com maior frequência. A moda da turma A foi de?…
> **Vários alunos:** 90.
> **Professora:** E da turma B?
> Vários alunos: 190.
> **Professora:** Pois é!? Que diferença! Numa turma a nota de maior frequência foi 90 e na outra foi de 190. Se não conhecêssemos os dados [apenas as modas]… [a] que conclusões chegaríamos?
> **Aluno:** Que a turma B é melhor.
> **Professora:** Com esses resultados da moda, será que ficam a achar que… que a turma B é um… máximo! Que a turma B parece bem melhor!... Mas é essa a realidade?
> **Vários alunos:** Não.
> **Professora:** Houve muitos 19 valores na B? Eu não sei… Digam lá…
> **Alguns alunos:** Houve 2… Só 2.

> **Professora:** [A moda] É a mais fácil de calcular, mas nem sempre nos diz… nos dá veracidade do conjunto de dados. O seu significado é mais limitado. (observação 3)

Nesta interação, Lia levanta alguma discussão sobre os valores da moda, inicialmente focada nos resultados obtidos e depois fazendo uma interpretação desses valores tendo em conta a observação dos restantes dados da amostra. A moda acaba por ser descrita como uma medida fácil de calcular, por se cingir à descrição de um aspeto muito específico da distribuição, atribuindo-lhe também um certo caráter limitativo.

Lia também apresentou a definição de classe modal: "é a classe com maior frequência" (observação 5) e mostrou aos alunos como identificar essa classe modal (de [65, 70[) numa situação com dados agrupados em classes, representados numa tabela de frequências e num histograma, *Consumo diário de leite em pó numa maternidade* (ver figura 57) que preparou e mostrou num *powerpoint* durante a aula. Adicionalmente, no histograma, a classe modal foi descrita por Lia como sendo a classe à qual corresponde a "barra mais alta". Lia também indicou o método para se obter uma estimativa do valor da moda, graficamente: "resulta de se procurar a abcissa do ponto de interseção dos segmentos marcados no histograma [da figura 57]" (observação 5).

Leite (g)	N.º de bebés
[45 , 50[	11
[50 , 55[	31
[55 , 60[	65
[60 , 65[	48
[65 , 70[	60
[70 , 75[	46
[75 , 80[	30
[80 , 85[	21

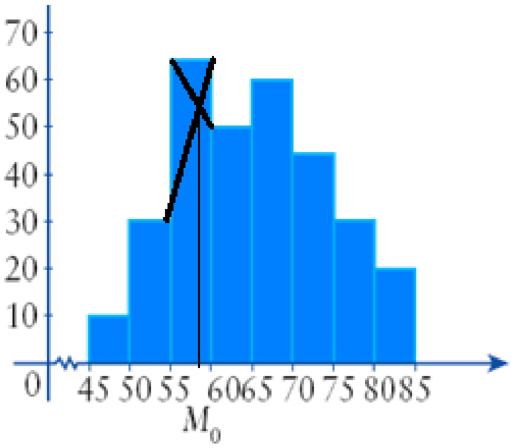


Figura 57: Representação dos dados num histograma (com indicação da classe modal).

Na correção da tarefa *Inquérito a 500 alunos sobre televisão* (t20), a moda é também identificada geometricamente através do histograma de frequências absolutas (figura 58).

Figura 58: Localização da moda no histograma.

Não foi indicada uma estimativa do valor da moda, nem a classe modal à luz dos dados e seu contexto. A classe modal foi, de certo modo, interpretada na t21 (*RSI por família)*. Na correção desta tarefa, Lia, ao pedir ao aluno Rui que lhe indicasse a classe modal e o "seu significado no contexto da situação apresentada, tal como era solicitado no enunciado", Lia obteve a seguinte resposta: "A classe modal é a de 200 para 300. Significa que há 34 famílias que recebem um rendimento entre 200 e 300" (observação 5). Para determinar geometricamente a moda, este aluno disse que bastava construir o histograma com as frequências absolutas e fazer a localização da moda na coluna da classe modal como se tinha feito na tarefa anterior, apontando para o histograma da t20 que estava ainda registado no quadro da sala de aula). Entretanto, também lhe mostrou o gráfico representado no seu caderno. Lia não deu seguimento à resolução da tarefa pelo facto de a aula estar a terminar.

Apesar deste aluno Rui não ter dito explicitamente que a classe modal correspondia à classe com maior frequência de famílias, foi possível observar que era isso que ele queria dizer, nomeadamente quando referiu como resolveria a questão da localização da moda graficamente.

Mediana e quartis

A mediana, o primeiro e terceiro quartil, e o diagrama de extremos e quartis foram também introduzidos através da tarefa *Qual a melhor turma, A ou B?* (t17) e referidos ao longo de diferentes tarefas, nomeadamente, t18, t19 e t20. Através das situações

Notas a Matemática no básico e *Notas a Inglês na escola*, Lia mostrou, em particular, como identificar os quartis numa tabela de frequências com as frequências relativas acumuladas, para dados não agregados e para dados agregados.

Tarefa *Qual a melhor turma, A ou B?* (t17) e Situação *Notas a Matemática no básico*

Quando solicita aos alunos para determinarem a mediana para as duas turmas englobadas na t17, Lia relembra que, depois de os dados estarem ordenados, a mediana corresponde ao "valor do meio ou valor central" (observação 3). Há um aluno que adverte a professora que pode não existir um valor central, no caso de haver um número par de dados. Lia explica que, nessa situação (que assinalou como sendo o caso da turma B), a mediana corresponde ao resultado da média aritmética dos dois valores centrais. Depois de apresentar, num *powerpoint*, os 23 dados da turma A, de forma ordenada, Lia questiona os alunos sobre o valor da mediana, respondendo, de imediato, alguns deles, que era 90:

> **Professora:** Então a mediana da turma A é 90… O que é que isso quer dizer?
> **Carlos:** Mais de 50% tirou negativa.
> **Professora:** Exato. Metade dos alunos tiraram 90 ou mais. Ficámos a saber como é a distribuição à direita e à esquerda da mediana… E para a turma B?
> **Alguns alunos:** A média aritmética dos valores centrais é 59,5 [mediana da turma B].
> **Professora:** O que é que isso quer dizer?
> **Eva:** Mais de metade da turma tem negativa.
> **Professora:** (…) 50% dos alunos teve nota inferior a 59.5. Já podem tirar alguma conclusão sobre a melhor turma? (…) Como professores, preferiam os resultados da turma A ou da B?
> **Vários alunos:** Da A.
> **Professora:** Mas a B tem 4 notas espetaculares! [4 dezanoves]… Na B, a mediana é muito mais baixa… (observação 3)

Quando questiona os alunos sobre o significado do resultado de mediana, as respostas dos alunos fazem uma alusão implícita aos 90 pontos [ou 9 valores de mediana], "mais de 50% tirou negativa" ou "mais de metade da turma tem negativa", tendo tido a professora o cuidado de ligar o valor do resultado à definição de mediana.

Com o cálculo da mediana e sua interpretação, os alunos já foram capazes de decidir sobre qual a melhor turma (A ou B) e optaram pela A. Lia ainda mencionou as quatro notas mais altas da turma B para tentar capturar indecisos. Como não obteve uma reação dos alunos, indicou também que a mediana da turma B era muito mais baixa do que a da turma A para mostrar, assim, a sua concordância com a resposta prévia dos alunos.

No prosseguimento desta tarefa, Lia introduziu os quartis. Apresentou-os como medidas de localização, através das quais é possível analisar a "maior ou menor concentração dos dados" entre intervalos cujos valores extremais são os dos quartis. Indicou que a mediana coincide com o segundo quartil e também que os quartis dividem a distribuição em quatro partes, cada uma compreendendo "aproximadamente 25% dos dados". Disse também que "A forma de calcular os valores dos quartis é análoga à forma como se localiza a mediana de um conjunto de dados" (observação 3). Mostrou aos alunos, na aula, como se determinava os quartis através de dois exemplos que incluíam duas amostras pequenas de dados quantitativos discretos. Na sua atuação seguinte, pediu-lhes para determinarem os quartis para cada conjunto de dados da tarefa *Qual a melhor turma, A ou B?* (t17) e, adicionalmente, para desenharem o diagrama de extremos e quartis, em paralelo, à semelhança do exemplo fornecido na tarefa. Lia reforçou que os diagramas de extremos e quartis, dispostos deste modo, facilitariam a realização de comparações.

Enquanto circulou pelos lugares dos alunos, Lia foi-se apercebendo que eles não estavam a saber calcular o primeiro e terceiro quartis, principalmente, por não saberem definir os subconjuntos de dados de modo a, a partir deles, determinarem os quartis pretendidos. Ao acompanhar esse trabalho dos alunos nos seus lugares, Lia foi ajudando e relembrando-lhes que "os valores à esquerda da mediana" permitiam identificar o primeiro quartil e "os valores à direita da mediana" (observação 3), o terceiro quartil. Também lhes deu indicações complementares de como representar essas informações num diagrama de extremos e quartis.

Ao pedir aos alunos comentários acerca dos resultados dos quartis para a situação da turma A e da turma B, obteve as seguintes respostas: O aluno Carlos disse "A turma B tem mais de metade de negativas! A turma A é a melhor…"; a Eva disse

> Quanto à mediana, sabemos que na turma A que 50% dos alunos tiraram igual ou mais de 90 pontos, na turma B tiraram igual ou mais de 59,5 pontos. Na turma A há mais concentração de dados entre 90 e 100 pontos e na turma B há mais concentração de dados entre 59,5 e 76,5 pontos". (observação 4)

Lia teve a seguinte interação com os alunos depois de ter escutado essas respostas:

> **Professora:** [Nas vossas respostas] Devem falar das medidas de localização e começar a interpretá-las. Qual é a mediana da turma A?
> **Vários Alunos:** 90.
> **Professora:** Isto quer dizer que 50 por cento da turma teve acima de 90 pontos. Qual é a mediana da turma B?
> **Vários alunos:** 59,5.
> **Professora:** Isto quer dizer que 50% da turma… eu já não digo para cima… Foi abaixo de 59,5 pontos, certo? Foram todos muito fracos?
> **Alguns alunos:** Não.
> **Professora:** Um ou outro. Há 4 boas notas e as outras são muito fracas. Aqui há uma diversidade maior… Há alunos muito bons e alunos muito fracos. E na turma A, existem alunos muito bons?... É uma turma mais homogénea ao nível dos conhecimentos. As suas notas são mais concentradas… entre 9 (mediana) e 10 (3.º quartil)… (observação 4)

Após esta interação, a professora disse aos alunos que eles poderiam alargar os seus comentários ao estabelecerem comparações entre amplitudes totais e intervalos interquartis da turma A e da B, cujos cálculos foram, de imediato, identificados por alunos e professora.

Neste excerto acima, depois de ter introduzido os quartis e ter pedido aos alunos para exporem os seus comentários face aos valores de quartis obtidos para cada turma, Lia voltou a confirmar a necessidade de os alunos estabelecerem ligações entre os conceitos e os resultados obtidos. Em termos de variabilidade dos dados, Lia assinalou a maior diversidade de notas na turma B e maior homogeneidade de notas na turma A, e também o intervalo com menor variabilidade entre os quartis Q2 e Q3 também para a turma A.

Lia questionou os alunos sobre como calcular a mediana num "conjunto muito grande de dados" indicando, por exemplo, os dados da situação *Notas a Matemática no básico* (observação 5), que foram mostrados no gráfico de barras (figura 59), havendo a seguinte interação:

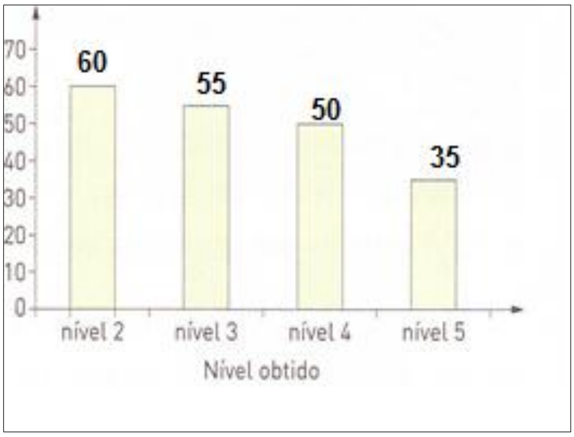


Figura 59: Diagrama de barras (notas a Matemática).

André: Por eliminação?
Professora: São muitos dados… vejam outra maneira…
Martim: Stora, e se fizermos o número total a dividir por dois? Dá 100. Vê-se mais ou menos onde vai estar… entre os 3 3 3 3….
Professora: Duzentos é par… Em que posição está a mediana? A mediana está entre o centésimo e o centésimo primeiro… É a média dos dois. Pelo gráfico, temos de ir até ao nível três para chegarmos a essas posições… E se vos perguntasse sobre o primeiro quartil?
Eva: 200/4.
Professora: Como fizeste?
Eva: Pela fórmula do livro…
Martim: O primeiro quartil deve estar entre o 50 e 51.
Professora: Entre o quinquagésimo e quinquagésimo primeiro… no nível 2 (observando o gráfico). Estão a entender?
Vários alunos: Não… (observação 5)

Nesta interação, Lia estava a conseguir obter as respostas desejadas. E desafiou os alunos a pensarem como chegar ao primeiro quartil, algo que não foi difícil para o aluno Martim. Depois de determinar entre que ordens estaria o valor da mediana e o do primeiro quartil, a identificação dos seus valores foi feita com recurso ao gráfico de barras pela professora. Os alunos que perceberam este processo de identificação da mediana e 1.º quartil experienciaram, também, desta maneira, uma forma de visualizar a mediana e 1.º quartil no gráfico.

Ao reparar que a maioria dos alunos não estava seguir o raciocínio do Martim nem o seu, para os ajudar, Lia sugeriu que todos se focassem mais nas frequências relativas acumuladas para identificação da mediana e quartis, e mostrou a tabela de frequências com os dados e quartis assinalados (figura 60), num *slide*, dizendo: “através da coluna das frequências relativas acumuladas, em percentagem, eu consigo facilmente identificar o 1.º, 2.º e 3.º quartil… Eles correspondem à localização dos primeiros 25%, 50% e 75% dos dados… como sabem" (observação 5). Contudo, relativamente a esta tabela (figura 14), não se teceram mais comentários nem explicações.

	Níveis	n_i	N_i	f_i (%)	F_i (%)	
$Q_1 = 2$ →	2	60	60	30	30	← Aqui 25%
$Q_2 = 3$ →	3	55	115	27,5	57,5	
$Q_3 = 4$ →	4	50	165	25	82,5	← Aqui 75%
	5	35	200	17,5	100	

Figura 60: Tabela de frequências exibida na aula.

Tarefa *Compara temperaturas máximas na Europa* (t18)

Alguns alunos, assim que viram e leram o enunciado desta tarefa, disseram não saber o que fazer para retirar conclusões, tal como a tarefa pedia. A professora apelou a que reparassem nas diferenças nos gráficos e que considerassem a "maior ou menor concentração dos dados" nos diferentes intervalos. Depois de dar alguns minutos para os alunos pensarem e escreverem as suas respostas nos seus cadernos, tem a seguinte interação:

Professora: Quem quer começar?
Mário: Eu!
Professora: Lê então.
Mário: Portugal continental obteve a maior temperatura do que a Europa. A temperatura mínima da Europa é inferior à de Portugal. (A professora pede ao Mário para ler mais alto)… Portugal continental obteve a maior temperatura do que a Europa.
Professora: Maior temperatura? A máxima das máximas… é o que queres dizer!...
Mário: [continua a leitura da sua resposta] A temperatura mínima da Europa é inferior à de Portugal. A amplitude interquartil na Europa é de 4 e a de Portugal é de 3. Portugal é mais quente do que a Europa.
Professora: Está bem…Quem fez mais conclusões? Diz lá, Martim (que levanta o braço).
Martim: 25% das temperaturas mais baixas da Europa… das mais altas da Europa [temperaturas máximas]… são inferiores à temperatura mais baixa… das mais altas de Portugal [temperaturas máximas]! (ouve-se vários risos)
Professora: O Martim conseguiu tirar uma conclusão! Perceberam o que ele disse?
Alguns alunos: Não…

> **Professora:** Ele disse que 25% das temperaturas mais baixas… da Europa… são mais baixas do que a mínima de Portugal. Mais? Ninguém mais?
> **Beto:** Oh stora… em Portugal, entre o terceiro quartil e o extremo máximo é mais concentrado.
> **Professora:** Outra coisa, em Portugal verificou-se que, entre o terceiro e o máximo, o espaço é pequenino; isto quer dizer o quê? Que os dados estão mais?...
> **Vários alunos:** Estão mais concentrados.
> **Professora:** Já na Europa, entre o terceiro quartil e máximo, os dados estão mais dispersos. (observação 4)

A professora não comentou a resposta do aluno Mário mas pareceu concordar com ela. Este aluno comparou os extremos inferiores e superiores de cada distribuição e ainda indicou a amplitude interquartil de cada uma delas. Lia também considerou que o aluno Martim deu um bom contributo quando observou que os primeiros 25% de dados da Europa eram inferiores à temperatura mais baixa registada em Portugal. Já o aluno Beto referiu-se ao intervalo mais estreito do diagrama de Portugal, entre Q3 e o máximo, como o "mais concentrado", situação que a professora também não deixou de realçar.

Lia tinha preparado um *slide* (figura 61) para mostrar alguns exemplos de conclusões que gostaria que os alunos se habituassem a fazer, usando os gráficos da tarefa 18, e procedeu assim na apresentação deste *slide*:

> Agora… vou mostrar-vos mais algumas conclusões que escrevi… O que vou mostrar não tem de ser igual para todos… O que interessa é que consigam olhar e consigam comparar… (lê as quatro frases do *slide* exibido e acrescenta) Podiam também dizer que, olhando para ali (apontando para os gráficos no *powerpoint*) que as temperaturas máximas de Portugal são sempre maiores, não é?... Comparativamente com as da Europa, certo?... Quer dizer que, se calhar… realmente as capitais distritais portuguesas são mais quentes… Podemos fazer este tipo de análise… (observação 4)

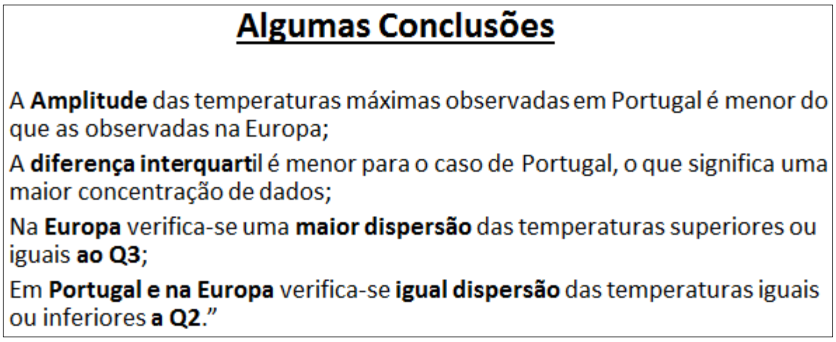


Figura 61: *Slide* apresentado na aula.

Nas quatro afirmações lidas pela professora, são comparadas as amplitudes, os intervalos interquartis, os primeiros 75% dos dados e os primeiros 50% dos dados entre as duas distribuições de temperaturas representadas em diagramas de extremos e quartis. A referência à variabilidade dos dados nesses intervalos foi feita através das palavras "dispersão" e "concentração". Nessas comparações, não são mencionados os valores concretos das medianas e dos quartis nem os valores das diferenças observadas. Mesmo depois de ter lido essas afirmações, Lia continuou mais um pouco a discussão à volta de comparações, falando mais do contexto dos dados e indo mais ao encontro do comentário feito pelo aluno Martim.

Tarefa *Associa entre conjuntos de dados discretos e diagramas de extremos e quartis* (t19)

Esta tarefa foi escolhida do manual e corrigida na aula. A professora pediu a duas alunas para irem ao quadro assentar os dados de cada horto ordenadamente e depois disse: "Só de olhar para os dados… que estão agora por ordem crescente… Será que não conseguem dizer qual é o diagrama de extremos e quartis correspondente a um e a outro?... Sem fazer os cálculos dos quartis!" (observação 4). Lia deparou-se com respostas diversificadas: houve alunos que lhe disseram que associavam o Horto *Linda Flor1* ao diagrama A, outros atribuíram esse horto ao diagrama B; houve também alunos que disseram que não sabiam como fazer.

Perante este cenário, Lia não questionou os alunos sobre as suas respostas e optou por pedir a dois alunos para irem ao quadro desenhar os diagramas, pedindo-lhes explicitamente para desenharem o diagrama A por baixo dos dados do Horto *Linda Flor1* e o diagrama B por baixo dos dados do Horto *Linda Flor2*. Houve alguns que reagiram, confessando que os diagramas, tal e qual como estavam a ser apresentados no quadro, não correspondiam. A professora prosseguiu, pedindo aos alunos para verificarem, novamente, se as associações estavam ou não bem feitas, com a seguinte sugestão que lhes deu: "Os dados entre o Q1 e Q2 estão muito mais concentrados no diagrama B do que no diagrama A". Depois dirigiu-se ao quadro e assinalou os quartis junto dos dados discretos e ordenados de cada distribuição, com a ajuda dos alunos. Alguns alunos intervieram na aula e voltaram a dizer que diagrama B correspondia aos dados do Horto *Linda Flor1* e o diagrama A ao Horto *Linda Flor2*. Lia teve a seguinte interação com os alunos de modo a perceber como é que eles chegaram às associações pretendidas:

Professora: Qual foi a razão da vossa escolha?
[Vários alunos falam em simultâneo, Lia tenta perceber o que o Beto diz]
Beto: No B a diferença é só 1 [referindo-se à distância entre Q1 e Q2] e no outro é mais.
Professora: [Assim que o Beto termina, Eva diz a sua resposta à professora, mas fala muito baixo e há algum barulho de fundo na sala]. Ouviram o que disse a Eva? A Rute já percebeu?
Rute: Acho que sim, Stora.
Professora: Ouviram o que a Eva estava a dizer?
Vários alunos: Não.
Professora: A Eva estava a dizer que o B é *Flor1* e que o A é *Flor 2*. Ela disse que [no B] a diferença entre Q1 e Q2 é 1 e entre Q2 e Q3 é 3,5. Diz lá, Eva…
Eva: E [no A] a diferença entre Q1 e Q2 é 3 e entre Q2 e Q3 é 3,5.
Professora: Porque é que também não podem dizer assim… Nesta parte aqui [do B]… há uma maior concentração de dados entre 9 e 10… Vamos ver quantos 9 e 10 temos aqui? [E conta 9 no total olhando para os dados originais ordenados]. No outro, entre 5 e 8, há cincos, seis, setes e oitos… Não existe tanta concentração de dados… Ali com os noves e dez havia muita repetição por isso é que na caixa… é mais estreita…
Beto: Stora… posso dizer uma coisa?
Professora: Diz…
Beto: Também do Q3… na *Flor 1*… ao extremo [superior]… a distância é maior…. Também os dados não se repetem muito…
Professora: De 11,5 para 5 é menor do que 13,5 para 20, certo?... Aqui o espaço teria se de ser maior… é uma outra observação… Mais alguma dúvida?
Martim: Stora, o que faz os quartis coincidirem?
Professora: Ah!... Vamos ver depois… [pausa grande] Isto quer dizer o quê? Se o Q1 coincide com o Q2 existe uma grande concentração de dados… E então, isto quer dizer o quê? Que os dados são iguais, é ou não é? Pensa…
Martim: Pois… (observação 4)

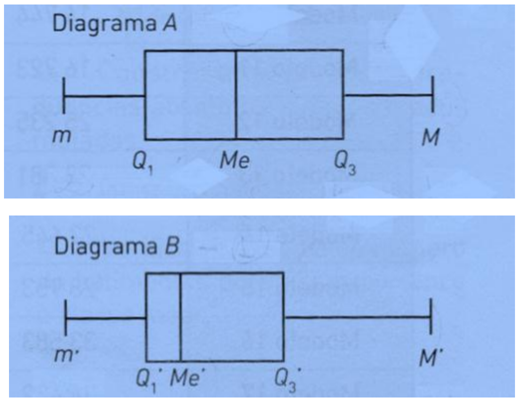


Figura 62: Diagramas de extremos e quartis (t19).

Apesar de a professora não ter questionado os alunos que estavam a estabelecer a correspondência entre dados e diagramas erroneamente, pensa-se que tal resposta se deve ao facto de os alunos se basearem só na mediana e não terem conhecimento que os diagramas da tarefa não estavam com a mesma escala, o que dificultava fazer a associação correta (ver figura 62). Contudo, a professora usou essa resposta errada e pediu voluntários para a registarem desta forma no quadro.

Durante a interação, as primeiras respostas dos alunos basearam-se na determinação dos três quartis para cada conjunto de dados, o que permitiu mencionar distâncias entre esses valores e associar corretamente o diagrama B aos dados horto *Linda Flor1* e o diagrama A aos dados do horto *Linda Flor2*. Para que os alunos não se ficassem apenas pelos cálculos, a professora pediu que observassem, juntamente com ela, os dados entre o Q1 e o Q2 em cada conjunto ordenado de dados também para que percebessem o motivo pelo qual a caixa do diagrama B era mais estreita entre Q1 e Q2. O aluno Beto acabou por usar o mesmo tipo de raciocínio para explicar porque é que o intervalo era maior entre Q3 e o valor máximo no diagrama B. No fim da interação, surgiu um pedido de esclarecimento do Martim: “O que é que faz os quartis coincidirem?”. A professora revelou-se um pouco surpreendida com essa questão, acabando por dar uma resposta que foi aceite pelo aluno que a questiono

Situação *Notas a Inglês na escola* e Tarefa *Inquérito a 500 alunos sobre televisão* (t20)

Para o caso de dados agrupados em classes, Lia referiu uma das formas mais rápidas de se localizar a mediana e quartis: “Através das frequências relativas acumuladas (…) e geometricamente no histograma de frequências relativas acumuladas eu só consigo dizer onde está cada quartil… Mas pela tabela identificámos apenas a classe mediana e as classes dos outros quartis” (observação 5). Apresentou, de seguida, a situação *Notas a Inglês na escola* através de uma tabela e de um gráfico com os quartis identificados (figuras 63 e 64):

Classificações (%)	n_i	N_i	f_i (%)	F_i (%)	
[0 , 25[	20	20	10	10	
[25 , 50[	120	140	60	70	← $Q_1 = Q_2$
[50 , 75[	40	180	20	90	← Q_3
[75 , 100]	20	200	10	100	

Figura 63: Tabela de frequências (com indicação da localização de cada quartil).

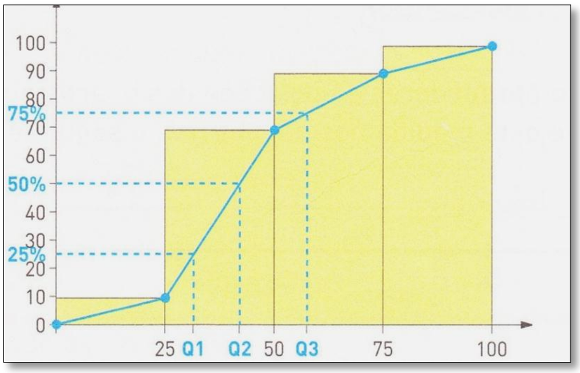


Figura 64: Histograma/Polígono de frequências
(com indicação da localização de cada quartil)

Ao mostrar esses dois gráficos (figuras 63 e 64), Lia destacou o facto de os Q1 e Q2 pertencerem à mesma classe, alertando que "isto pode acontecer … e porquê?… isto quer dizer que os dados deveriam realmente estar muito concentrados aí, e vocês podem verificar que sim (apontando para a tabela de frequências)" (observação 5).

Lia explicou aos alunos que os quartis poderiam ser identificados numa tabela de frequências, através das suas classes, e que a única forma de se obter um valor aproximado de cada quartil, com dados agrupados em classes, seria graficamente com o histograma e polígono de frequências relativas acumuladas. Através das tabelas, era possível identificar a classe mediana e a dos restantes quartis e tomar-se a marca da classe como uma estimativa da mediana, o que não foi sugerido por Lia.

Ainda, baseando-se nesta situação, Lia explicou que era possível desenhar um diagrama de extremos e quartis por baixo da dupla representação histograma/polígono, aproveitando-se as cinco informações que o eixo dos *xx* deste histograma fornecia, nomeadamente: mínimo, máximo e os três quartis.

Os alunos tiveram oportunidade de construir um diagrama de extremos e quartis desta forma na tarefa *Inquérito a 500 alunos sobre televisão* (t20). Na figura 65 apresenta-se o trabalho realizado por um aluno no quadro para corrigir a última alínea desta tarefa:

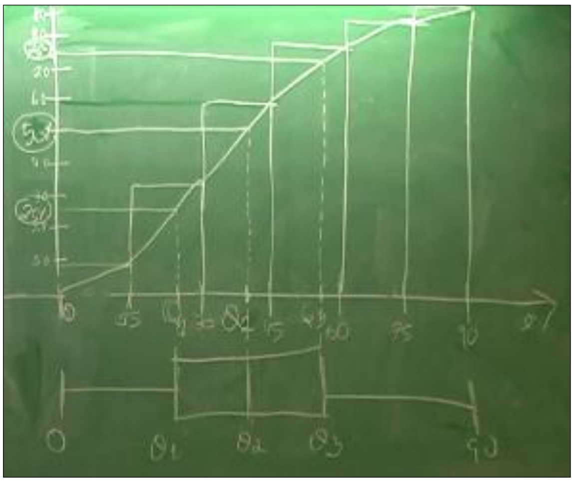


Figura 65: Ligações entre representações (histograma/polígono de frequências relativas acumuladas e diagrama de extremos e quartis).

Na correção desta alínea, não houve comentários a essas representações obtidas, nem aos quartis, nem uma análise global à forma da distribuição. Houve, sobretudo, a preocupação de se alcançar, corretamente, as representações, tal e qual, como pedido no enunciado. É também de constatar que, na construção e na apresentação do diagrama de extremos e quartis, não se procurou indicar uma estimativa numérica para cada quartil.

6.3.3.2 Reflexões sobre o tópico na aula

As medidas de localização são, na perspetiva de Lia, acessíveis para a maioria dos alunos, que devem saber como determiná-las. Nas aprendizagens realizadas pelos alunos neste tópico, Lia considera que, quanto à média e mediana, há uma maior tendência para fazerem mais confusão e de trocarem uma com a outra. Contudo, depois de os alunos as aprenderem e praticarem ficam com a noção que as três medidas de localização "são valores intermédios entre o mínimo e o máximo". Na sua opinião, a noção intuitiva que os alunos retêm da moda é "aquilo que aparece mais vezes… o que está mais na moda!"; a noção de média está associada ao seu algoritmo e a de mediana diz respeito à ordenação dos dados e à procura do valor central ou a fazer a média dos dois valores centrais nestes dados.

Ainda, ao comentar sobre as ideias que os alunos deveriam possuir destas três medidas, Lia classificou-as quanto à sua representatividade: "deviam reter que a média dá um número que, muitas vezes, não é representativo (do conjunto), é preciso atender à variação (…) a média [das notas de um aluno] pode corresponder a uma nota que o aluno não merece" (entrevista 2). Lia tem a conceção que os seus alunos ficaram com a ideia que, destas três medidas, a mediana talvez seja a mais representativa por ser "menos sensível a valores extremos" e "por ficar ali no meio, ali localizada, não dá mais informação" (entrevista 2). Notoriamente, para Lia, a mediana é, comparativamente à média, a medida menos influenciável e, por isso, mais apta a representar um conjunto de dados, embora conferindo-lhe uma certa limitação por não prestar esclarecimentos sobre os dados além da sua localização central.

A professora é também da opinião que, das três medidas de localização, média, moda e mediana, a média é a que acaba por ficar mais presente para os alunos e é tendencialmente a mais usada por eles na escola e fora dela. Para contrariar essa tendência, e para que os alunos revisitassem de uma forma mais aprofundada esses conceitos, Lia teve algum cuidado em preparar tarefas com dois conjuntos de dados (t17, t23), que permitissem que as médias fossem próximas para que os alunos fossem levados a procurar outras medidas, para melhor argumentarem as suas conclusões. Lia considera que algumas tarefas propostas na aula foram, de algum modo, inspiradas pela formação que frequentou sobre Estatística (juntamente com a investigadora), nas quais colocou "perguntas diferentes", pedindo que os alunos tirassem mais conclusões, além dos cálculos habituais:

> Acho que fui um pouco influenciada pela ação de formação, aqueles diálogos, com maior ou menor orientação do professor, saíam de lá com os miúdos a chegar às conclusões, [os alunos] eles é que faziam o trabalho todo… Aqui em Portugal os professores é que fazem o trabalho todo… Tenho essa sensação. Os [nossos] alunos não estão habituados a essas perguntas mesmo que se as colocasse fora do tema da Estatística… Não estão habituados a concluir… a ir mais além… a questionar os dados… (entrevista 2)

A tarefa *Qual a melhor turma, A ou B?* (t17) foi preparada pela professora já com a intenção de provocar mais alguma discussão à volta das medidas de localização e suas utilidades. A professora desejava que os alunos sentissem necessidade de utilizar outras medidas além da média, moda e mediana, nomeadamente, os quartis e o diagrama de extremos e quartis para poderem responder à questão proposta na tarefa de uma forma mais fundamentada:

> **Professora:** O objetivo desta questão era que os alunos chegassem à conclusão de que as medidas de localização, pelo menos, média, moda e mediana, não iam servir para tirar conclusões… Não iam ser grande garantia para eles tirarem conclusões acerca disso… A média era praticamente a mesma (…) A moda da turma B era 19, acho eu, a moda da turma A era 9. Há muitos 19. Se não tivesse aqui os dados… e dissesse, a moda de uma turma é 19 e a moda da outra turma é 9. A que é que chegavam à conclusão?... São medidas que podem ajudar mas que, em alguns casos, nalgumas distribuições, não nos ajudam grande coisa, não é?… Não são representativas das distribuições… Então, ao fazer o diagrama e ao introduzir os quartis, e quando eles fizeram realmente os diagramas, eles deram muito diferentes e deram mais concentração aqui nos dados… nos 9 e nos 10… dos 25% de dados que estão aqui… E começaram a ver… Se calhar tiveram mais noção da distribuição de cada uma das turmas… uma noção real. O objetivo desta tarefa era mesmo esse.
> **Investigadora:** E quanto às conclusões a que eles chegaram?... Eles viram os diagramas de extremos e quartis no enunciado da tarefa…
> **Professora:** Eles até perguntaram se [os diagramas] tinham de dar sempre assim [como no exemplo fornecido], lembras-te?
> **Investigadora:** Sim, lembro-me…
> **Professora:** Eles fizeram e depois criticaram os gráficos. Depois tiveram noção quanto menor algum quadradinho [dentro da caixa de bigodes] mais concentração de dados estão ali. (…) Eles leram as [suas] conclusões… falaram da média, moda e mediana e chegaram à conclusão que uma turma [A] era melhor do que a outra apesar da moda ser 19 [turma B]. Os resultados aqui [turma B] foram piores.
> **Investigadora:** Era a esse resultado que querias que eles chegassem?
> **Professora:** Sim… Também proporcionar contactos com outras medidas… quais as medidas mais úteis. Eles aperceberem-se da realidade das distribuições, não é? Porque nem sempre a média… a média é muito sensível a valores extremos, a moda é uma medida muito fácil de calcular mas nem sempre também é representativa. A mediana talvez seja a mais representativa… porque fica ali no meio… mas também não sabemos o que se passa para a direita e o que se passa para a esquerda… é um bocado isso…, não é? (entrevista 2)

Este excerto mais longo revela os objetivos que a professora tinha ao propor a tarefa aos alunos, e mostra também definições informais que a professora detém dos conceitos referidos. Constata-se o facto peculiar de a professora analisar a utilidade das medidas de localização como representantes do agregado, que foram também algumas das ideias intuitivas que desejou transmitir aos alunos.

Lia também comentou que detetou algumas dificuldades nos alunos quando construíram os diagramas de extremos e quartis nessa tarefa 17, que tinham de ser apresentados em paralelo nos mesmos eixos cartesianos. Lia recordou que os alunos não se estavam a lembrar onde é que os quartis estavam localizados no diagrama:

> Acho que notei dificuldades em construírem o diagrama de extremos e quartis. Principalmente então quando tinham de colocar um em cima do outro. O tempo era pouco. Já não tinham bem noção do que fazer… [risos]… Já não sabiam onde é que estava o primeiro quartil… de uma aula para outra. Eles precisavam de mais tempo… Os exercícios práticos são necessários para os alunos terem aptidão para fazerem mais alguma coisa. (entrevista 2)

Apesar de a tarefa *Qual das gémeas a melhor aluna?* (t23) ter sido proposta fundamentalmente para se introduzir o conceito de desvio-padrão, Lia fez algumas reflexões em torno das medidas de localização calculadas nas suas primeiras alíneas:

> Ao fim ao cabo… esta tarefa é idêntica à anterior que comentei [Tarefa *Qual a melhor turma, A ou B?* (t17)]. Porque lá está… peço para calcular a média, moda e mediana… depois dava tudo exatamente a mesma coisa [treze valores]. Então, se eles não tivessem a distribuição [os valores dos dados], mas se soubessem que média, moda e mediana é igual, então se calhar também diriam que era tudo igual [mesmo os dados]… [esta tarefa permitiu] terem a noção das três medidas… do uso delas… Quando perguntava se a média dava um bom indicador, lá está… Acho que não tiveram grande dificuldade em responder e, mesmo na pergunta acerca dos desempenhos das duas alunas, se eram ou não eram semelhantes (…) (entrevista 2)

Mais uma vez, nesta tarefa (t17), os alunos eram explicitamente incitados a questionar os resultados obtidos nos cálculos das medidas de localização dado que seriam todos iguais. Lia também queria que indagassem até que ponto se deveriam basear nessas medidas para decidirem qual das gémeas era a melhor aluna.

Lia considera que a tarefa *Compara temperaturas máximas na Europa* (t18) que propôs na aula seria um pouco mais desafiante do que a t17, pois os alunos não teriam de construir os diagramas, apenas comentá-los. Quando falou desta tarefa, a professora disse já estar à espera de ter alunos que lhe iam perguntar que tipo de resposta é que teriam de escrever. Lia considerava que os alunos não estavam habituados a ser confrontados com este tipo de questões, por isso, achava normal que fosse receber várias questões ou que viesse a notar uma certa inatividade na aula. Comentou, da seguinte forma, as dificuldades que emergiram no desenvolvimento desta tarefa:

> A dificuldade maior quando apresentei aqueles dois diagramas de extremos e quartis [da t18]… falei do que é que eles representavam… e só coloquei uma

> questão, que era basicamente ‘o que é que conclui?’ Portanto, não coloquei alíneas. Eles sentem-se perdidos. Se tivesse feito umas alíneas… Qual o mínimo ou o máximo? Qual é o primeiro quartil? Onde se concentram mais os dados? Onde se concentram menos? Onde estão os últimos 25% de dados? Eles… provavelmente não tinham dificuldades em responder. Já quando se pede ‘representa um gráfico’ e para concluir… eles têm muitas dificuldades em fazer um texto sobre o assunto e tirar conclusões… Têm aquela preguiça porque não estão habituados a este tipo de abordagem. Estão habituados a uma abordagem mais direta até porque os exames nacionais… na Matemática as abordagens são sempre todas diretas. Habitualmente não são usadas perguntas… de questionamento… Depois não há tempo… não é num exame que se vai classificar esse tipo de abordagem... (entrevista 2)

Lia considera que, se a tarefa, em vez de pedir uma conclusão, tivesse várias questões que indicassem especificamente o que deveria ser observado na comparação dos gráficos, os alunos, de uma maneira geral, saberiam o que fazer em cada uma delas. A tarefa, nos moldes em que estava constituída, era desafiante e fora do habitual na disciplina de Matemática.

Como também não estava à espera que os alunos fossem assinalar vários comentários na sua resolução, Lia trouxe algumas frases escritas que apresentou num *powerpoint* para que os alunos ficassem com um conjunto mais alargado de comentários. Quanto à interação que houve no momento de correção da tarefa, Lia disse:

> Foi aqui que realmente constatei que alguns tiraram algumas conclusões… Eu também assinalei algumas, mas alguns até tiraram conclusões diferentes das minhas, embora algumas delas mais simples. O Martim dizia que não conseguia tirar nenhuma e até conseguiu… Mas, se calhar, até foi quem tirou a melhor conclusão. (entrevista 2)

Apresentou-se a Lia um conjunto de respostas hipotéticas de alunos à análise de um diagrama de extremos e quartis da Situação *Respostas hipotéticas sobre um diagrama de extremos e quartis* (anexo 2), que correspondiam à distribuição das notas de uma turma de 24 alunos no último teste realizado de uma disciplina. Ao ler essas respostas hipotéticas, Lia disse de imediato quanto à primeira: “Não acho que a turma seja razoável, com 50% de notas positivas e 50% de notas negativas”; não fez comentários à segunda afirmação e achou que a terceira e primeira eram semelhantes: “E a terceira é um pouco a repetição da primeira…”; quanto à última afirmação, disse que: “Na última poderia estar [o valor] 12 em vez de terceiro quartil” (entrevista 3).

A professora fez comentários adequados a estas afirmações, ao avaliá-las globalmente, Lia considerou-as comuns junto dos alunos, e reparou também que não houve nenhum aluno a fazer referência à “amplitude interquartil”. Considerou ainda que este tipo de tarefa não tem o mesmo potencial, para a aprendizagem, que uma tarefa com dois conjuntos de dados, para colocar os alunos a fazer comparações: “Muito sinceramente… seria muito mais produtivo se houvesse outra turma para se poder fazer comparações entre elas… senão é tudo muito local…” (entrevista 3).

Lia pediu aos alunos para trazerem a resolução da tarefa *Reação positiva a um medicamento* (t22), que foi proposta para trabalho de casa, numa folha à parte. Lia A professora comentou as resoluções juntamente com a investigadora. Nesta tarefa, a professora considerou que a questão mais desafiante era a última alínea, que pedia para fazer a análise comparativa dos diagramas. Contudo, também achava que, tendo os alunos já lidado com algumas tarefas que envolviam a construção e interpretação de diagramas de extremos e quartis, eles não teriam, de uma maneira geral, muitas dificuldades em representar os dados desta forma e apresentar, por escrito, “algumas conclusões”. A maioria dos alunos apresentou os cálculos das medidas de localização para os três conjuntos de dados pedidos no enunciado. Lia detetou algumas incongruências nas respostas a esses valores nalguns alunos, embora a indicação do procedimento fosse correta (no momento, foi impossível determinar exatamente o erro em cada situação, pelo motivo de os conjuntos de dados serem de 21, 29 e 50 elementos). Ao analisar as respostas dos alunos, no que diz respeito à alínea da análise comparativa, Lia chegou à conclusão que não eram “muito completas”, apontando algumas lacunas, que se apresentam sintetizadas a seguir: alguns alunos indicaram os valores da amplitude de cada distribuição, valores das medianas e intervalos de menor ou maior dispersão, mas não fizeram a comparação pedida entre homens e mulheres; os alunos que identificaram quais os intervalos da caixa com maior ou menor concentração não fizeram essa mesma análise para a distribuição global.

Lia também considerou que as descrições feitas pelos alunos acabaram por não estabelecer a ponte entre a representação construída (diagrama de extremos e quartis) e o contexto do enunciado da tarefa:

> Eles aqui… por exemplo,… nem sequer falam em termos reais envolvendo o contexto, restringem-se muitos às percentagens… valores para cima e para

> baixo… Não se libertam muito nas conclusões, não é? … Se eu não visse o enunciado não sabia do que estavam a falar… Eles não desenvolveram muito as conclusões… Até poderiam fazer alguma afirmação… falta aí um bocado de tirar conclusões para além das percentagens, dos números… (…) e relacionar com o contexto… Lá está… também não estão habituados! (entrevista 2)

Lia colocou apenas uma questão sobre Estatística no teste final do 3.º período, a questão *Compara o aproveitamento em três disciplinas* (t29). Ao analisar as respostas dos alunos juntamente com a investigadora, Lia considerou que eles conseguiram identificar a disciplina com melhor aproveitamento (Biologia), apesar de considerar que, ao nível das justificações, há respostas diferentes: umas com algumas incongruências (em que os alunos não assinalaram os intervalos certos para o que queriam afirmar), umas com mais informação do que outras, outras em que a é realizada a comparação entre partes de diagramas, e noutras em que não. Ao ler a resposta do aluno Martim, Lia destacou-a como uma "resposta correta e rápida" (entrevista 2): "A disciplina foi a de Biologia, pois através do gráfico é possível saber que 75% das classificações em Biologia eram maiores ou iguais a 10, o que não acontecia com as outras disciplinas, pois a Matemática o valor da positiva é inferior a 75% e a Inglês é inferior a 50%".

Apesar de confessar que esta resposta poderia estar escrita de uma forma mais clara, menciona que a entende, explicando que a decisão do aluno Martim se baseou na "percentagem de positivas em cada disciplina" (entrevista 2) observada nos diagramas de extremos e quartis. Ou seja, esta percentagem, na Biologia, era de cerca de 75%, na Matemática, era inferior a 75%, e no Inglês, era inferior a 50%.

6.3.3.3 Discussão do tópico

CETE. Foram propostas na aula tarefas de origem diversa. Umas foram propositadamente produzidas pela professora; outras foram selecionadas de manuais, que envolviam a comparação de conjuntos de dados através das medidas de localização. Ao longo da sua resolução, os conceitos foram sendo introduzidos e essas tarefas acabaram por ser corrigidas pela professora em interação com os alunos. Nesses momentos, Lia usou, muitas vezes, um discurso que visava a interligação das medidas com a observação da variabilidade dos dados, para se poder tirar conclusões sobre os mesmos dados. Realçou que conclusões sobre os dados através dessas medidas estatísticas deveriam ser feitas com muita cautela porque eram medidas que podiam ocultar, de certa forma, a variabilidade dos dados. Essa preocupação da professora não foi transmitida no contexto de dados agregados em classes. Lia, em situações de maior variabilidade nos dados, tomou a mediana como uma medida mais representativa do que a média e também do que a moda. Lia também apresentou algumas situações e tarefas que tinham o propósito de identificar valores de quartis em representações tabulares e gráficas. Nestes contextos, acabou por ser mais trabalhada a construção dessas representações, o cálculo/identificação das medidas, e a identificação de um intervalo de variação de dados onde houvesse maior concentração de dados, do que o interpretar e comentar essas representações e cálculos à luz da problemática subjacente.

A visualização da mediana e quartis graficamente num diagrama de barras e histograma surgiu na aula de uma maneira desunida, por exemplo, sem destacar caraterísticas da forma da distribuição global dos dados, o que parece não ter contribuído para o aprofundamento destes conceitos. Neste tópico, a professora não achou que fosse frutífera a utilização das funcionalidades estatísticas da calculadora gráfica (até porque, na maioria das tarefas, os alunos tiveram que lidar com conjuntos pequenos de dados), optando por aproveitar as situações propostas nas aulas para que os alunos consolidassem, com lápis e papel, os cálculos das medidas de localização e fizessem algumas representações (diagrama de extremos e quartis e histograma). É de destacar que a opção de elaborar tarefas que pedissem simultaneamente os valores das medidas de localização e também que esses valores fossem analisados e comparados de uma forma crítica, em vez de pedir, apenas, o cálculo de medidas, sobretudo da média, terá sido afetada, simultaneamente, pela sua experiência docente e como formanda.

CA. À medida que a professora foi propondo as tarefas e foi observando como é que os alunos lidavam com elas, foi-se apercebendo de algumas facilidades e também dificuldades que encontraram. Lia detetou dificuldades na determinação dos quartis, em particular, do Q1 e Q3, e também na interpretação da tabela de frequências para dados agrupados em classes, de modo a retirar-se a informação correta para os cálculos a fazer. Apesar de a professora considerar que a interpretação dos resultados ou do diagrama de extremos e quartis não eram tarefas fáceis para os alunos, achou que estes, de uma maneira geral, foram progredindo na aquisição da capacidade de fazer observações, de comentar os resultados, dentro das oportunidades que lhes foram dadas na aula.

Lia optou por esclarecer prontamente os alunos quando estes forneciam respostas erradas ou com partes de respostas desadequadas; raramente avançou para os questionar sobre como tinham alcançado essas respostas. Houve também respostas corretas dadas pelos alunos às quais não pediu justificações. De uma maneira geral, neste tópico, houve um acesso limitado aos raciocínios dos alunos. Lia, ao ser confrontada com respostas de alunos nalgumas tarefas que propôs para trabalho de casa e também com as respostas dadas a uma questão num teste, constatou também dificuldades, na maioria dos alunos, em expressar, por escrito, as comparações pedidas entre diagramas de extremos e quartis.

CE. A professora expressa claramente as definições dos conceitos e menciona alguns aspetos particulares de alguns deles. Comunica também a necessidade de se criticar os resultados através da sua interpretação, tendo em conta o contexto dos dados, nalgumas tarefas propostas. Na aula, Lia, ao analisar os resultados das medidas de localização face à variabilidade presente nos dados (em situações que tinham o objetivo de alertar os alunos que, por trás de duas médias iguais ou próximas, podiam existir conjuntos de dados muitos diferentes), pode também ter induzido os alunos em erro, nomeadamente, ficando a pensar que a média não é uma medida representativa, em distribuições simétricas de variabilidade reduzida, ou até com alguma variabilidade, revelando algumas dificuldades em raciocinar sobre variabilidade e média, em simultâneo.

CC. Lia desenvolveu tarefas para a sua prática que tinham como objetivo fazer com que se questionasse mais os assuntos envolvidos relacionando-os com os dados e contexto de cada situação, algo que se propôs trazer na sala de aula influenciada pela sua participação na ação de formação. Este facto terá contribuído para que ela própria ampliasse a sua visão do currículo escolar e ganhasse alguma perceção do desafio para o professor e para o aluno em lidar com tarefas que exigem mais comunicação e mais discussão dos resultados.

6.3.4 Medidas de dispersão

When comparing groups or making inferences, we need to look at center and spread together: the signal and the noise around the signal.
Garfield e Ben-Zvi (2008, p. 204)

A análise que se faz do conhecimento didático em Estatística de Lia, relativa a este tópico, tal como refletido na sua prática, divide-se em três secções. Na primeira (desvio-padrão), apresenta-se uma contextualização das medidas de dispersão e analisa-se como é que essas medidas foram desenvolvidas na prática de Lia; na segunda secção (reflexões sobre o tópico na aula), apresenta-se algumas reflexões da professora sobre o ensino e aprendizagem do tópico. Na terceira (discussão do tópico), realiza-se uma discussão em torno das *Medidas de dispersão*, guiada pelas quatro dimensões do modelo do conhecimento do professor em Estatística.

6.3.4.1 Desvio-padrão

Amplitude total e amplitude interquartil são medidas de dispersão identificadas por várias vezes no tópico *Medidas de localização* (tópico anterior), nomeadamente, nas tarefas que pediam a determinação de quartis, construção de diagramas de extremos e quartis e análises a esses diagramas. Neste tópico, Lia deu principal destaque ao desvio-padrão como medida de dispersão, pelo que isto vai-se refletir ao longo da descrição e análise de dados que se apresenta de seguida.

Tarefa *Qual das duas gémeas a melhor aluna?* (t23)

Foi através desta tarefa que Lia introduziu as medidas de dispersão variância/desvio-padrão e também classificou, pela primeira vez, amplitude total e amplitude interquartil como medidas de dispersão: “Vamos recorrer às medidas de dispersão amplitudes totais e interquartis e variância e desvio-padrão… que é a raiz quadrada da variância… para podermos tirar conclusões sobre os dados” (observação 6). Ao aperceber-se que os alunos já tinham obtido os resultados da amplitude total e da amplitude interquartil para as notas de cada gémea, sem recorrerem ao *menu STAT* da calculadora gráfica, Lia interage do seguinte modo com eles:

> **Professora:** Já podemos fazer comparações… O que podemos concluir com essas medidas tão simples?
> **Mário:** No intervalo de amplitude 2 [amplitude interquartil] estão 50% dos resultados…
> **Professora:** Então, o que é que concluis com isso?
> **Mário:** São mais concentrados na Joana [amplitude interquartil é 2] do que na Mariana [amplitude interquartil é 10].
> **Professora:** Ah!!... Como o intervalo é mais pequeno… há mais concentração de dados ali, certo? Então, [nesse mesmo intervalo] há menor quê? (nenhum aluno responde)… Vamos estudar a concentração/dispersão dos dados! (…) Há uma coisa que quero que vocês tenham muita atenção! Essas medidas… principalmente a amplitude [total] é a mais fácil de calcular mas só nos dá informação sobre os valores extremos. Eu sei que existe maior dispersão mas os valores [dos dados] do meio, não sei como são… podem ser 1, 2, 17 e o resto 13 [valores]… Então, para ver se realmente existe maior desvio à média, que é 13, ou não… vamos calcular medidas de dispersão mais atualizadas… para verificar a variabilidade dos dados através da variância e do desvio-padrão… Ao preencherem a tabela e seguirem os passos na ficha vamos chegar ao valor da variância e do desvio… (observação 6)

Nesta intervenção, Lia apercebe-se que os alunos não tiveram dificuldades na determinação das medidas amplitude total e amplitude interquartil. E, centrando-se na amplitude total, descreve-a como uma medida simples que dá indicações sobre os extremos e maior ou menor concentração/dispersão dos dados no intervalo, mas que não permite saber como é efetivamente a distribuição entre os seus extremos. Neste sentido, introduz a variância e o desvio-padrão como medidas que permitem analisar a distribuição quanto à variabilidade dos dados.

De seguida, apresenta-lhes um quadro (quadro 14) que terá de ser preenchido (exemplifica como se completa o quadro para a nota dez) e também indica que o quadro compõe cálculos auxiliares úteis para determinar-se os valores da variância e do desvio-padrão.

Quadro 14: Quadro apresentado na aula para apoiar o cálculo do desvio-padrão

Notas Joana	Freq. Abs.	Desvio	Quadrado dos desvios	Desvio de cada valor x_i relativamente à média $\bar{x}$
x_i	n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 n_i$
10	1	10-13= -3	$(-3)^2 = 9$	$9 \times 1 = 9$
12				
(…)				
				Soma Total (*)

Enquanto circulava pela sala de aulas, foi tirando dúvidas quanto ao preenchimento do quadro pedido. Em equipa, professora junto ao quadro e alunos no lugar, seguindo os passos assinalados na ficha de trabalho para se obter o valor da variância e do desvio-padrão, alcançaram um valor de variância de σ^2 =20/7 (cujo numerador resultou da soma dos valores da última coluna do quadro) e de desvio-padrão $\sigma = 1{,}69$ com os dados da gémea Joana (figura 66):

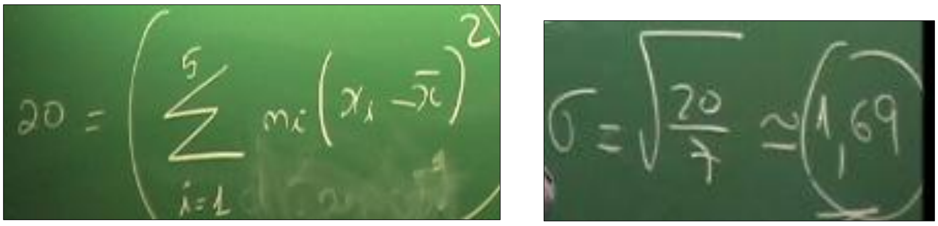


Figura 66: Resultados do numerador da fórmula da variância e do desvio-padrão.

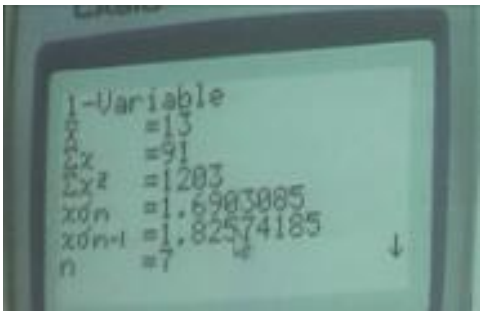
O excerto seguinte mostra qual o sentido que foi dado ao resultado dos desvios-padrão calculados, depois de a professora ter dado tempo aos alunos para determinarem no lugar o desvio-padrão para a gémea Mariana:

> **Professora:** O que é que o desvio-padrão nos diz? O que mede?
> **Martim:** A variabilidade.
> **Professora:** Em relação a quê?... A variabilidade do conjunto de dados relativamente à… média. O quão perto ou longe se está da média [13].
> **Professora:** Então digam lá, quanto deu o da Mariana? (ouve-se respostas diversificadas) 0,64 não pode ser!... Quanto deu?
> **Eva:** 4,2.
> **Professora:** E com duas casas decimais?
> **Eva:** 4, 17 (outros alunos dizem concordar com o valor).
> **Professora:** Vamos comparar os dois valores. Qual é o maior? O que concluo sobre o desvio-padrão?
> **Vários alunos:** O da Mariana é maior…
> **Luzia:** As notas da Mariana afastam-se mais… (Há outros alunos a falar em simultâneo)
> **Professora:** Exato. Mariana tem notas mais afastadas da média. O desvio-padrão é uma medida que indica a grande ou pouca variabilidade dos dados relativamente à média. Perceberam? Agora, precisam de praticar a sua fórmula. (observação 6)

Lia, quando ouviu a resposta de 0,64 de desvio-padrão para as notas da Mariana, considerou-a logo como uma resposta errada, embora não tenha explicado o motivo nem determinado onde estaria o erro junto do aluno. Muito provavelmente, considerou que este aluno errou nalgum cálculo auxiliar ou no cálculo final deste desvio-padrão.

Assim que ouviu o valor correto, preocupou-se em comentar este valor quando à sua variabilidade relativamente à média. Nesta interação, emerge uma ideia informal acerca do desvio-padrão, apresentado como uma medida que relaciona a variabilidade dos dados com a média. Na comparação dos dois valores, o mais elevado é descrito como o mais "afastado" da média. Desta forma, Lia explicou aos alunos como devem exprimir-se relativamente ao desvio-padrão.

Para ensinar os alunos a chegar ao resultado do valor do desvio-padrão através das calculadoras gráficas, Lia começa por exibir na tela o *software* da calculadora gráfica *Casio* e depois faz o mesmo para o da *Texas Instruments*. Os passos essenciais que mostrou nas duas calculadoras consistiram na inserção de listas de dados no menu Estatística da calculadora gráfica (lista das notas e lista das frequências de cada nota, colocando ordenadamente as notas da aluna Joana), no uso do *menu CALC* com respeito a uma variável estatística (*1 VAR* ou *1Var Stat*) associando-o às listas criadas e, com isso, na determinação simultânea de um conjunto de medidas estatísticas. Lia assinalou o resultado em $x\sigma n$ ou em σx de, aproximadamente, 1.69 como sendo o resultado procurado para o desvio-padrão nas duas calculadoras (figura 67):



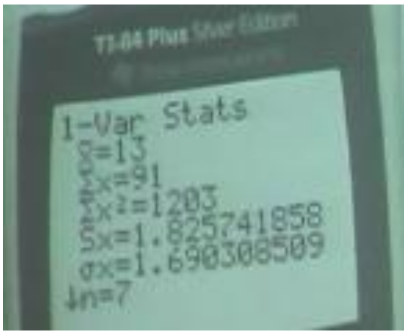


Figura 67: Valor do desvio-padrão das notas da Joana nas duas calculadoras.

Lia revelou domínio nas duas calculadoras gráficas à medida que as dificuldades dos alunos foram surgindo. A professora detetou algumas dificuldades na obtenção das medidas estatísticas nas duas calculadoras. A maioria dos alunos que possuía *Casio* não conseguiu obter os resultados que a professora exibiu na tela. Para colmatar essa dificuldade, a professora pediu-lhes para confirmarem se as listas estariam bem assinaladas na calculadora gráfica, alertando-os para irem ao *menu EXE* e no *SET* atribuir as listas de dados previamente criadas, a *Xlist* e a *Freq*. Já com a calculadora

Texas Instruments, houve alguns alunos que não obtiveram os cálculos que estavam assinalados no quadro por se terem esquecido de colocar as duas listas à frente da instrução *1-VARS STAT*, tendo Lia sugerido que voltassem a colocar as instruções seguidas das listas usadas.

Tarefa *Média e desvio-padrão em dados agrupados* (t24)

Lia leu, como habitualmente fazia, o enunciado desta tarefa em voz alta, para a turma, e explicou que a primeira alínea – determinação da marca da classe (que tinha sido abordada no tópico anterior) – iria ajudar a calcular a média e o desvio-padrão. Também registou a fórmula do desvio-padrão no quadro e indicou que *ni* correspondia à frequência e que cada *xi* teria de corresponder à marca de cada classe.

Depois de dar tempo para os alunos trabalharem na tarefa, pediu dois voluntários para irem ao quadro corrigi-la. Como nenhum aluno se ofereceu, pediu diretamente ao aluno Luís para fazer a parte correspondente à turma A e à Rute para fazer a parte da turma B. Cada um destes alunos ocupou uma parte do quadro e registou a média e o desvio-padrão.

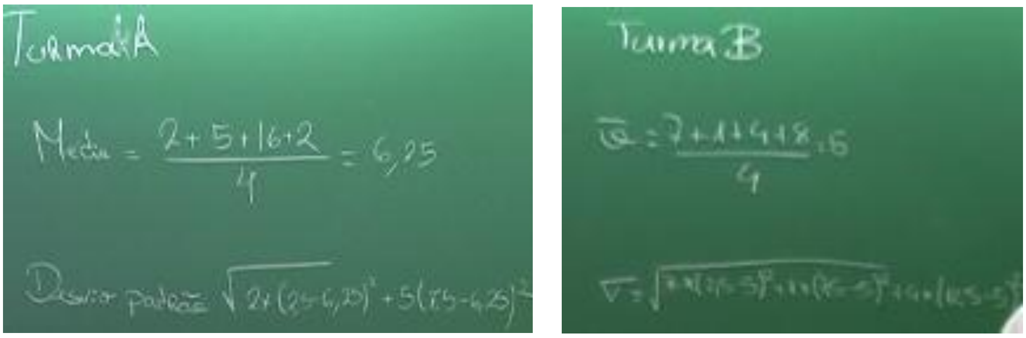


Figura 68: Médias das duas turmas.

Depois de observar o que os alunos estavam a fazer, já quando tinham iniciado o cálculo do desvio-padrão, a professora interpelou a turma e os alunos que estavam no quadro acerca dos resultados das médias obtidas (figura 68):

> **Professora:** Toda a gente concorda com os resultados da média [que estão no quadro]?
> **Alguns alunos:** Não…

> **Professora:** Quem não concorda? (Alguns alunos dizem que não concordam. Lia, colocando a sua atenção sobre Telma, disse) Telma, como farias?
> **Telma:** Tem de ser a marca da classe a multiplicar pelo seu *ni*.
> **Professora:** Exato. Então o que representa o *ni*?
> **Alguns alunos:** A frequência.
> **Professora:** Então, só há 4 alunos? (observação 7)

Lia notou que Luís e Rute tinham os dois feito exatamente o mesmo erro na média: somaram as quatro frequências e dividiram por quatro, em vez de procederem à divisão entre as somas de cada classe multiplicada pela sua frequência e o número total de notas. Também se observou que estes dois alunos, quando usaram a fórmula do desvio-padrão, tinham substituído corretamente *ni* pela frequência e *xi* pela marca da classe. Esta situação, embora no contexto de dados agrupados, parece demonstrar que os alunos sabem em que consiste o procedimento da média, mas não interpretam corretamente os dados de modo a retirar a informação que necessitam para fazer o cálculo.

Lia pede a Telma para ir ao quadro corrigir as médias registadas pelos seus dois colegas. Luís esperou que Telma corrigisse a sua parte e Rute, como tinha percebido qual era o erro, corrigiu a sua parte. Luís e Rute também retificaram esses valores na fórmula do desvio-padrão, nas respetivas situações. Enquanto circulava na aula, e depois destes reparos, Lia encontrou alguns alunos com dificuldades na interpretação da tabela que estava no enunciado (com duas colunas: classificação e *ni* (frequências)) e no cálculo do valor da média. Houve a seguinte interação entre a professora e a turma depois de os valores da média e do desvio-padrão estarem corretamente calculados e registados no quadro:

> **Professora:** Os cálculos já estão feitos. Vamos olhar para os resultados. Que é que eu quero? Quero que comparem as médias…
> **Luzia:** Elas são quase iguais. [Turma A- média de 11.1; Turma B- média de 10.75]
> **Professora:** Mas isto quererá dizer que as distribuições são iguais?
> **Alguns alunos:** Não…
> **Professora:** Então qual é a medida estatística que nos dá essa informação sobre… eventuais diferenças?
> **Luzia:** O desvio-padrão.
> **Professora:** [Ele] Na turma A foi de?
> **Vários alunos:** 3.61
> **Professora:** E na turma B?
> **Professora e alunos:** 6, 57.
> **Professora:** O que é que isso significa?
> **Luís:** Há mais dispersão de dados na turma B.

> **Professora:** Há mais dispersão em torno da média! Embora as médias sejam quase iguais, o desvio-padrão é diferente, isso indica que as distribuições são diferentes. Há alguma dúvida? Sabem como se calcula o desvio-padrão na calculadora? (observação 7)

Média e desvio-padrão surgiram interligados na abordagem de Lia. Pelo facto de as médias terem dado um valor aproximado, Lia questionou os alunos acerca das distribuições, procurando saber se, nessas circunstâncias, as distribuições também seriam similares. A professora procurou que os alunos resolvessem essa questão através da comparação dos valores dos desvios-padrão. Embora houvesse um aluno a referir que a dispersão no conjunto B era maior, foi a professora que acabou por mostrar como é que os alunos deveriam explicar-se. Assim, o desvio-padrão foi considerado uma medida útil para indicar que as distribuições podem ser diferentes e para descrever a maior ou a menor variabilidade presente nos dados face à média. No excerto, não houve uma referência explícita às classificações de cada turma, mas foi feita uma descrição global das distribuições das duas turmas relativamente às suas variabilidades como interpretação do valor do desvio-padrão, sem recurso a qualquer representação gráfica.

Depois desta interação, Lia também explicou as instruções para a determinação do desvio-padrão na calculadora gráfica *Texas Instruments* para o caso da turma A. Deixou também claro que se deve usar a calculadora gráfica para calcular o desvio-padrão e todas as outras medidas estudadas até ao momento, particularmente, quando o conjunto de dados for grande.

6.3.4.2 Reflexões sobre o tópico na aula

Lia confessa que, nos momentos de preparação deste tópico para a sala de aula, não fez um "estudo exaustivo ao manual" (entrevista 2), sobretudo por falta de tempo. Considera que, mesmo assim, no manual adotado, não encontrou o tipo de tarefa que desejava para iniciar o tópico, optando por criar tarefas (com distribuições de dados e perguntas específicas, como foi o caso da t23) que englobassem questões que interligassem os conteúdos da forma como pretendia, ou seja, que proporcionassem mais alguma discussão acerca dos resultados, não ficando somente pelos cálculos, que é algo que encontra com mais facilidade nos manuais:

> Se calhar no manual não encontrei duas distribuições que tivessem a mesma média, moda e mediana. Eu tinha um objetivo diferente. As tarefas do manual pedem: calcula a média, a moda e a mediana… [no manual] não encontrei o que queria… Eu queria fazer essas perguntas (apontando para as questões e para a informação incluída na ficha de trabalho que integrava a tarefa *Qual das gémeas é a melhor aluna?* (t23), que solicitava em, particular, a comparação dos valores dos desvios-padrão de duas distribuições e que se retirasse conclusões)… Se calhar, com a moda, média e mediana sendo iguais… esses valores não são suficientes para tirar-se uma análise real… com conclusões… seria preciso outros valores [nomeadamente, desvio-padrão]… mas há aqui também tarefas interessantes (indicando, como exemplo, a tarefa t25 do manual adotado)… Não tive grande tempo… (entrevista 2)

Lia indicou, a título de exemplo, uma tarefa do manual, inserida no tópico das medidas de dispersão (t25) que poderia ter proposto na aula caso tivesse tempo. Efetivamente, na aula, Lia não propôs nenhuma tarefa do manual, apenas as que criou e incluiu numa ficha de trabalho, como foi o caso das tarefas t23 e t24, especialmente preparadas para o ensino do tópico das medidas de dispersão. Nesta ficha de trabalho, Lia inseriu adicionalmente a definição de desvio-padrão, a fórmula da variância e do desvio-padrão e cálculos auxiliares para se chegar mais rapidamente, com recurso a lápis e papel, ao valor do desvio-padrão.

Apesar de Lia se considerar uma “adepta das tecnologias”, tomando-as como uma estratégia de ensino importante para um maior entendimento dos conteúdos e que também poderia servir para cativar o aluno, acabou por lhes dar pouco uso no ensino deste tema. Na sua perspetiva, a aprendizagem de qualquer tema de Matemática exige fundamentalmente “trabalho manual” e a tecnologia (calculadora gráfica, computador) deve ser usada de uma forma bastante ponderada. Ao explicar porque é que no tema da Estatística a usou pela primeira vez no cálculo do valor do desvio-padrão, Lia refere:

> Para que é que preciso de usar as listas da calculadora gráfica para calcular uma média com poucos dados?... Agora, o desvio-padrão tem uma fórmula mais complicada… demora mais tempo a calcular (mesmo que seja a fazer os cálculos parciais na calculadora). Acho muito bem que se use as funcionalidades do *menu* Estatística para o calcular… desde que [os alunos] entendam o que é pedido. (entrevista 2)

De facto, na tarefa introdutória do desvio-padrão (*Qual das duas gémeas a melhor aluna?*, t23), foram usadas as funcionalidades do menu Estatística da calculadora gráfica, mas só depois de a professora ter apresentado a fórmula do desvio-padrão para este ser calculado no caderno. Lia teceu alguns comentários ao ter insistido no uso da fórmula do desvio-padrão na aula:

> Aqui, para calcular as medidas de dispersão achei que deveria orientar sem ser na calculadora gráfica, achei que eles deveriam ter um contacto com a fórmula [do desvio-padrão]… e depois saberem como fazer nas calculadoras (colocavam as listas e determinavam o desvio-padrão)… E manualmente para terem noção de que a variância, mais propriamente, o desvio-padrão vai, neste caso, medir a variabilidade de um conjunto de dados relativamente à média. Achei que deveria orientá-los no cálculo [através dos passos que indicou na tarefa]… O desvio-padrão numa delas devia ser muito menor do que no outro. Existe mais variabilidade na Mariana do que na Joana, porque o valor do desvio-padrão era maior… O desvio-padrão era mais uma das medidas que [os alunos] poderiam recorrer para verificar e fazer um estudo quanto à variabilidade dos dados relativamente à média. (entrevista 2)

Para Lia, era importante que os alunos tivessem contacto com a calculadora gráfica para aprenderem a calcular o desvio-padrão, com o menu de Estatística da calculadora, mas optou por fazer isso só depois de os alunos terem experimentado a fórmula do desvio-padrão através das indicações que deu na ficha de trabalho. Na sua opinião, esta atividade iria ajudá-los na compreensão da noção de desvio-padrão, como medida que mede a variabilidade dos dados em relação à média, ideia que está expressa na fórmula e seu cálculo.

A investigadora pediu à professora para ler e comentar sobre a situação *Desvio-padrão elevado* (anexo 2) que lhe mostrou. A sua leitura suscitou alguns comentários. Lia afirmou que, para qualquer situação indicada, tudo "dependeria da distribuição de dados" mas "a segunda é um pouco confusa" por abordar um gráfico. Na primeira situação, não é mencionada a média, pelo que Lia levantou a questão: "Aqui haverá variabilidade relativamente à média?... Não sei bem?!". Ainda assim consideraria como respostas adequadas a terceira e a quarta, ainda que, para a quarta, acrescentou tratar-se de uma situação muito específica e que "haver uma maior variabilidade com alguns *outliers*" (entrevista 3) também seria uma resposta válida.

Com estes comentários, Lia denotou possuir algumas ideias informais acerca do conceito de desvio-padrão, que surgiu associado à média e à medição da variabilidade dos dados, e revelou compreender a influência que um ou mais *outliers* podem ter no resultado no desvio-padrão.

Quando a investigadora questionou Lia sobre a noção que os alunos deveriam reter acerca do desvio-padrão, esta afirmou, prontamente, "é mesmo a ideia da variabilidade… o desvio-padrão corresponde a um valor que indica a variabilidade que existe dos dados relativamente à média da distribuição" (entrevista 2). Assim sendo, Lia revelou, mais uma vez, ter uma noção de desvio-padrão associada à medição da variabilidade dos dados face à média, considerando que foi essa mesma noção que tentou privilegiar nas aprendizagens dos seus alunos.

6.3.4.3 Discussão do tópico

CETE. Na aula, propôs fundamentalmente um trabalho à volta de duas tarefas que criou para o ensino deste tópico. Se tivesse mais tempo para lecionar o tópico, teria proposto mais tarefas de consolidação, nomeadamente, recorrendo às do manual escolar. Os dados e questões incluídos nessas tarefas propostas foram propositadamente selecionados pela professora para gerarem mais discussão sobre os resultados. Cada tarefa comportava dois conjuntos de dados, os quais iriam produzir medidas de localização próximas (nomeadamente, médias) que iriam conduzir à necessidade de se calcular uma medida de variabilidade, de certo modo, mais informante do que amplitude total e amplitude interquartil, o desvio-padrão, para assim se poder tirar mais algumas ilações na comparação dos resultados. Na opinião de Lia, a opção de calcular o desvio-padrão manualmente, com apoio da fórmula, que expressa diferenças numéricas entre dados e média, "desvios", poderia facilitar o entendimento dos alunos do desvio-padrão como uma medida de variabilidade face à média. Lia ensinou a usar a calculadora gráfica para calcular o desvio-padrão, mas recomendou sobretudo o seu uso para conjuntos de dados grandes. Mesmo assim, na aula, não houve oportunidade de se trabalhar com conjuntos grandes de dados. O desvio-padrão foi referido como uma medida que pode evidenciar diferenças em distribuições mas não recorreu a representações gráficas para sustentar essa ideia intuitiva que lançou.

CA. Apesar de os alunos terem trabalhado neste tópico somente com duas tarefas, Lia foi detetando algumas dificuldades dos alunos no cálculo manual do desvio-padrão e também através das calculadoras gráficas. Lia reparou no desafio sentido por alguns alunos, ao nível da leitura dos dados agregados em classes em tabelas de frequências, em que selecionaram dados errados da tabela para o cálculo da média. Esta situação iria, por acréscimo, influenciar um cálculo errado do desvio-padrão. Os alunos, nos seus comentários comparativos relativamente aos valores de desvio-padrão obtidos, foram fazendo algumas referências a esses valores, relacionando-os com variabilidade, mas Lia complementou sempre as suas respostas, tendo o cuidado de destacar que deveriam entender o desvio-padrão como uma medida que mede a variabilidade dos dados de cada conjunto relativamente à sua média

CE. Lia indicou a fórmula da variância/desvio-padrão e expressou algumas ideias informais acerca dos conceitos amplitude total, amplitude interquartil e variância/desvio-padrão. Lia mostrou dominar a calculadora gráfica quando ajudou alunos a superar dificuldades neste tópico. Numa entrevista, Lia demonstrou saber da pouca resistência do desvio-padrão face a *outliers*; contudo, não abordou este assunto na aula.

CC. Através das tarefas propostas e desenvolvidas na aula, foi possível observar que Lia deu principal destaque à determinação de medidas estatísticas de dispersão e comparação de resultados. Apontou, também, algumas limitações dessas medidas para retirar-se conclusões sobre os dados e seu contexto. Também deu alguma ênfase à complementaridade entre medidas de localização e medidas de dispersão na análise de dados.

6.3.5 Dados bivariados e regressão linear

> *Reasoning about covariation commonly involves translation processes among raw numerical data, graphical representations, and verbal statements about statistical covariation and causal association.*
> J. Moritz (2004, p. 228)

Este tópico é analisado em quatro secções. As duas primeiras (correlação e regressão linear, e modelo linear) incidem sobre noções centrais do tópico e explicam o modo como essas noções emergiram e foram desenvolvidas na prática da professora. A terceira secção (reflexões sobre o tópico na aula), apresenta algumas reflexões de Lia acerca do ensino e aprendizagem deste tópico; a quarta secção (discussão do tópico), expõe uma discussão do tópico guiada pelas quatro dimensões do modelo do conhecimento didático do professor em Estatística.

6.3.5.1 Correlação e regressão linear

Nesta secção foram definidas quatro partes (Dados bivariados; Coeficiente de correlação linear; Centro de gravidade; e Reta de regressão) que indiciam, em particular, a sequência e enquadramento das noções associadas à correlação e regressão linear tomada por Lia na sua prática. Ao longo desta secção, Lia usou o recurso multimédia *escola virtual* para apresentar vários exemplos e situações que envolviam dados bivariados. Na utilização deste recurso, sobressaíram algumas descrições e explicações que forneceu relativamente às relações bivariadas, ao uso do coeficiente de correlação linear, ao centro de gravidade e à reta de regressão que serão descritas e analisadas de seguida.

Dados bivariados

Lia realçou alguns exemplos de conjuntos de duas variáveis estatísticas cujo estudo da existência de relação entre ambas parecia ser relevante (por exemplo, com dados sobre criminalidade e taxa de desemprego; áreas de casas e seus valores de mercado; e notas de Matemática e de Inglês de uma turma) mas não lhes explicou as razões da sua opinião, nem perguntou aos alunos as suas opiniões. Na representação destes dados, destacou o diagrama de dispersão (ou nuvem de pontos) com os dados bivariados representados sob a forma de pares ordenados. Ao abordar o diagrama de dispersão, indicou que o estudo sobre dados bivariados tinha em vista analisar a "relação entre as duas variáveis". Este diagrama foi referido pela professora como uma representação particularmente "útil para realçar propriedades entre os dados" e a "forma da distribuição". A observação dos dados representados neste diagrama foi referida como um ponto de partida fundamental para a análise da eventual existência de relação entre as variáveis, sendo classificada como "não existe", "correlação positiva" ou "correlação negativa" (observação 7).

Uma das situações apresentadas pela professora, do recurso *escola virtual*, foi *Notas de Inglês versus Notas de Matemática* (figura 69). Sobre esta situação, concluiu não existir uma relação entre notas de Inglês e de Matemática de uma turma, explicando que, no diagrama de dispersão, os dados se apresentavam "de uma forma aleatória" (observação 7).

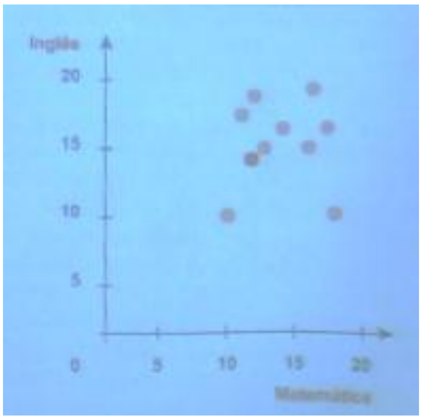


Figura 69: Nuvem de pontos (*Notas de Inglês* versus *Notas de Matemática*).

Situação *Nuvem de pontos das temperaturas de capitais europeias*

Nesta situação, a professora solicitou aos alunos a análise da relação entre as variáveis temperatura mínima (eixo dos *xx*) e temperatura máxima (eixo dos *yy*) de várias capitais europeias, representadas na nuvem de pontos seguinte (figura 70). Lia questionou os alunos sobre a existência de uma relação entre essas variáveis:

> **Professora:** Será que existe alguma relação? Esta nuvem de pontos estava como a outra?! (referindo-se ao *Exemplo Notas de Inglês versus Notas de Matemática*) com uma forma aleatória?
> **Vários alunos:** Não.
> **Professora:** Será que existe relação?… Olhem para os valores [no diagrama]... O que é que acontece quando a temperatura mínima aumenta?… a outra vai aumentando?
> **Vários alunos:** Aumenta.
> **Professora:** Então… A temperatura máxima também vai aumentando. Há aqui uma relação entre elas. Quando uma aumenta, a outra também aumenta, não de uma forma proporcional, não é?... Mas vai aumentando. A correlação… Há correlação positiva entre as duas variáveis, pois quando uma aumenta a outra também aumenta. Ok? […] A nuvem de pontos é ascendente… (observação 7)

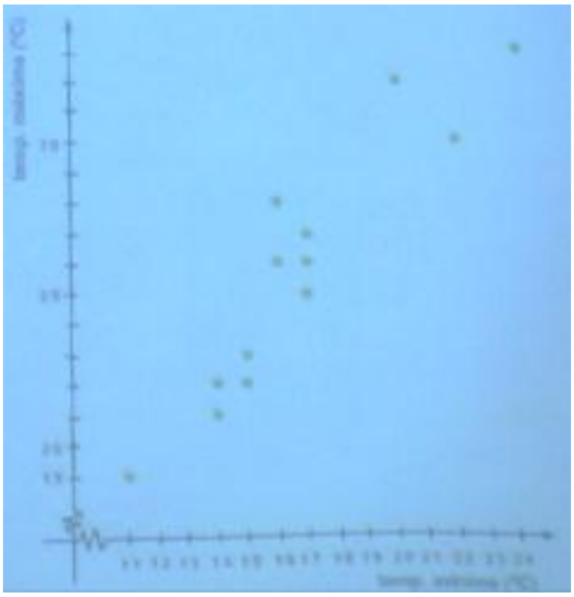


Figura 70: Nuvem de pontos de temperaturas (mínima e máxima) de capitais europeias.

Nesta interação, Lia começou por pedir aos alunos para comparem a nuvem de pontos que tinham visto anteriormente (com os dados distribuídos de uma forma aleatória) com a da situação atual. A maioria dos alunos indicou que as distribuições em causa eram diferentes. Lia fez uma leitura da nuvem de pontos fazendo referência à tendência global de crescimento conjunto que observou nas variáveis: à medida que os valores de uma aumentam, os da outra também aumentam. Lia estabeleceu uma associação entre a tendência global de crescimento que observou nos dados representados e a existência de uma correlação positiva. Apesar de ter também especificado que o crescimento observado não podia ser traduzido de uma forma proporcional, não fez comentários quanto à existência de alguns dados da distribuição que não seguiam essa tendência, e os alunos também não a interpelaram sobre esses aspetos.

Coeficiente de correlação linear

Ao selecionar e analisar vários exemplos da *escola virtual* de diagramas de dispersão com distribuições diversificadas, a professora foi introduzindo o coeficiente de correlação linear como uma medida que toma valores numéricos entre -1 e 1 (inclusive) e classificou-o quanto ao sinal (positivo ou negativo) e à intensidade da correlação (perfeita, forte, moderada ou fraca). Associou também o sinal do coeficiente de correlação linear ao declive da tendência global dos dados observados no diagrama de dispersão/nuvem de pontos. Apontou também a maior proximidade dos dados entre si com uma reta específica, a de regressão, quando o valor do coeficiente linear, em termos absolutos, era próximo de 1 ou igual a 1. Ao sintetizar essa informação, na aula, referiu ainda:

> Quando *r*=1, quando há correlação linear perfeita positiva, implica que os pontos estão mesmo alinhados sobre uma reta… o declive da reta é positivo (…) Imaginemos agora que nós temos um coeficiente de correlação *r*= - 0.9. O que é que vai acontecer?... Os pontos não vão estar todos sobre uma reta, mas vão estar todos juntinhos… Vai haver uma reta próxima deles todos. Dizemos [neste caso] que a correlação é forte, mas negativa. Quanto à intensidade da correlação, temos também a correlação moderada… mas quanto mais o valor de *r* [coeficiente de correlação] se aproxima de zero ou a correlação é fraca ou não existe (…) À reta que se aproxima muito dos pontos nós vamos chamar reta de regressão. Ela tem uma equação. Se colocarmos os dados na calculadora [é possível], a calculadora dá-nos uma equação do tipo *y*=a*x*+b… Dá-nos o valor de a, o valor de b… os seus valores aproximados… e o nosso coeficiente de correlação… que é o meu *r*. (observação 7)

De seguida, mostra um vídeo que indica os procedimentos a tomar numa calculadora gráfica para se obter a reta de regressão e respetivo coeficiente de correlação em simultâneo. Neste excerto transcrito da aula, Lia começou por referir-se ao coeficiente de correlação linear. Ao atribuiu-lhe alguns valores numéricos, explicou como é que esses valores poderiam ser traduzidos quanto ao modo como os dados poderiam estar posicionados na nuvem de pontos face a uma reta (de regressão) e também expôs quando é que poderia existir uma reta (de regressão) próxima desses dados bivariados. Lia deixou transparecer, também, o que entende por reta de regressão e expressou a utilidade da calculadora gráfica na determinação da expressão analítica dessa reta, bem como na obtenção do valor do coeficiente de correlação linear.

Lia propôs algumas situações da *escola virtual* e tarefas (por exemplo, t25) que solicitavam que se fizesse a correspondência entre diagramas de dispersão de diversas distribuições e valores de coeficientes de correlação também variados. Nessas situações e tarefas, os alunos forneceram as respostas corretas relativas às associações estabelecidas e, de uma maneira geral, os comentários justificativos foram sucintos e baseados no sinal (positivo ou negativo) da tendência global dos dados e na intensidade da correlação observada (classificada habitualmente como: nula, fraca, moderada, forte ou perfeita). Para comentar a intensidade da correlação, alguns alunos apenas se referiram à concentração ou dispersão dos dados entre si nos diagramas. Essas distribuições foram sempre caracterizadas de uma forma geral e também não foi observada qualquer ideia que envolvesse a reta de regressão. Por exemplo, ao questionar os alunos sobre o possível coeficiente de correlação que corresponderia ao diagrama de dispersão da figura 71, Lia teve a seguinte intervenção:

> **Professora:** Aqui há dúvidas? Qual é o coeficiente de correlação [que associam ao diagrama]?
> **Vários alunos:** É zero!
> **Professora:** Mas porquê? Vamos lá, digam?
> **Miguel:** Não há…
> **Professora:** Porquê?
> **Rafael:** Há dispersão.
> **Professora:** Porque os pontos estão muito… muito dispersos. De facto, não existe uma nuvem ascendente… nem descendente. (observação 8)

Nesta interação, observa-se que os alunos e a professora são muito sintéticos nas suas respostas. O aluno Miguel não deu uma resposta completa e, quando o Rafael referiu existir dispersão no gráfico, a professora aproveitou para fazer um comentário de modo a suportar a resposta desse aluno, sem o questionar sobre a sua resposta.

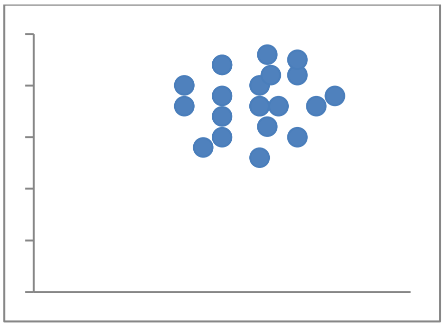


Figura 71: Diagrama de dispersão ao qual foi correspondido um valor de coeficiente de correlação zero.

Centro de gravidade

Lia salientou da *escola virtual* um diagrama de dispersão que tinha a particularidade de ter o centro de gravidade marcado. Definiu-o como o ponto de abcissa média da variável x e ordenada média da variável y. Esse diagrama de dispersão que exibiu na aula incluía, também, um eixo horizontal e outro vertical centrado no referido centro de gravidade. Na interação que teve com os alunos acerca desse gráfico, Lia destacou a localização privilegiada do centro de gravidade no diagrama de dispersão (figura 72), na análise da tendência global dos dados e na confirmação do sinal da correlação:

> **Professora:** Quando a correlação é negativa a maior parte dos pontos vai estar neste e neste [apontando para os quadrantes pares formados com os eixos centrados no centro de gravidade da distribuição que estavam incluídos no diagrama de dispersão] mas não são todos, ok? E se a correlação for positiva?
> **Ivone:** [aponta para os quadrantes ímpares].
> **Professora:** Exato. Este centro de gravidade pode ajudar-nos a verificar se a correlação é positiva ou negativa… Mas olhando para o diagrama de dispersão vocês devem conseguir ver se vai haver correlação positiva ou não, ok? (observação 7)

Nesta interação, Lia associou a correlação negativa com a tendência descendente dos dados observada ao longo dos quadrantes pares marcados sobre o centro de gravidade no diagrama de dispersão. Nesta situação específica, todos os dados estavam inseridos nesses quadrantes; contudo, a professora, na sua explicação, teve o cuidado de referir que, no caso de haver correlação, a maior parte dos dados estariam nessas circunstâncias, podendo existir exceções. Não emergiram outras explicações que ajudassem a perceber, por exemplo, a estreita relação entre o centro de gravidade e a reta de regressão ou até do centro de gravidade com o valor do coeficiente de correlação linear. Durante o tema da Estatística, esta foi a única vez que a professora mencionou o centro de gravidade de uma distribuição bidimensional na aula.

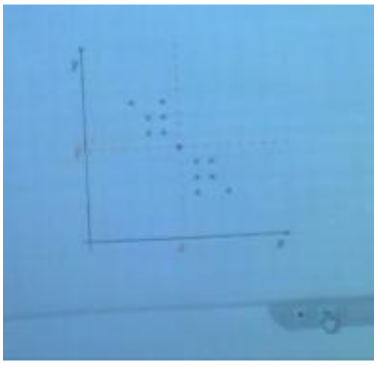


Figura 72: Centro de gravidade destacado no gráfico de dispersão.

Reta de regressão

À medida que Lia foi gradualmente introduzindo os conceitos de coeficiente de correlação linear e de centro de gravidade, realizou também uma abordagem intuitiva acerca da reta de regressão. Essa reta foi descrita pela professora como aquela "que melhor se aproxima de todos os pontos do diagrama de dispersão" (observação 7). A exposição dos conteúdos, com o apoio da *escola virtual*, terminou quando Lia referiu a utilidade da reta de regressão para se "estimar ou prever valores da variável dependente ou da independente, dentro e fora dos intervalos de variação de cada uma delas" (observação 7), desde que os valores usados e determinados fizessem sentido no contexto dos dados. Nesta situação, Lia não deu a conhecer que, dada uma reta de regressão, esta deveria ser essencialmente usada para estimar valores da sua variável dependente. Adicionalmente, a professora ainda referiu, de passagem, que a reta de regressão era também conhecida como "reta dos mínimos quadrados", lendo uma afirmação a esse respeito do *powerpoint* que exibiu:

> Há muitas retas que se podem desenhar, mas um dos critérios mais comuns para definir essa reta é o de tornar mínima a soma dos quadrados dos desvios dos pontos em relação à reta. A essa reta chama-se reta de regressão ou reta dos mínimos quadrados. (observação 7)

Não houve nenhuma situação nesta aula que permitisse perceber se os alunos estariam a entender o critério que a professora lhes leu sobre a definição de reta de regressão. Os termos "desvios" e "quadrados dos desvios" já tinham sido usados em aulas anteriores na designação de parcelas da fórmula do desvio-padrão. No entanto, não surgiram explicações nem comentários sobre a estrutura dessa reta na aula, tendo apenas Lia relembrado que já lhes tinha mostrado como determinar essa reta na calculadora gráfica juntamente com o valor do coeficiente de correlação.

Apesar de Lia não ter relacionado explicitamente a reta de regressão com o centro de gravidade (que já tinha sido objeto de algumas considerações na aula observada 7) exibiu um diagrama de dispersão que, além dos pontos marcados, incluía a sua respetiva reta de regressão e o centro gravidade $(\bar{x}, \bar{y})$. Este último surgia assinalado sobre a reta de regressão (figura 73).

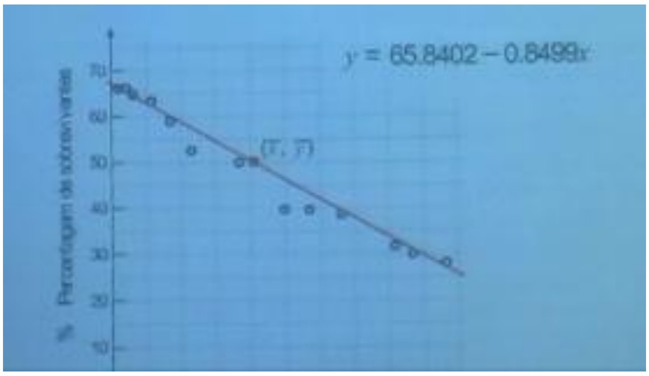


Figura 73: Nuvem de pontos com reta de regressão e centro de gravidade desenhados.

6.3.5.2 O modelo de regressão linear

Apresenta-se de seguida uma descrição e análise do modo como a reta de regressão foi usada como modelo linear através das tarefas propostas na aula, nomeadamente, nas tarefas *A evolução do crescimento de Pedro* (t27) e *Distribuição do número de médicos e farmácias por regiões* (t28).

Tarefa *A evolução do crescimento do Pedro* (t27)

Lia selecionou esta tarefa do manual escolar. Na sua apresentação aos alunos disse-lhes que teriam de analisar a relação entre as duas variáveis em estudo (*idade* e *altura de uma pessoa*) e recorrer à reta de regressão para estimar um valor para cada uma das variáveis. Informou, também, que teriam de utilizar a calculadora gráfica para representar esses dados num diagrama de dispersão e determinar os resultados pedidos.

Na resolução dessa tarefa, Lia pretendia recorrer às duas calculadoras gráficas (de marcas diferentes) mais usadas na aula. Preparou o computador e projetou uma delas, começando pela da marca *Casio*. Pediu um voluntário para trabalhar no computador a sua resolução. O aluno que se voluntariou, Luís, inseriu os dados nas listas da calculadora gráfica projetada e, com a ajuda da professora que lhe relembrou os passos a seguir na calculadora, mostrou rapidamente o diagrama de dispersão com a reta de regressão assinalada. Não foi feito nenhum comentário à visualização obtida. Curiosamente, o enunciado da tarefa não pedia o valor do coeficiente de correlação mas este valor surgiu em simultâneo com a equação da reta de regressão na calculadora gráfica. A interação de Lia com os alunos centrou-se nos valores obtidos para a equação da reta de regressão e no valor do coeficiente de correlação (que também podem ser observados na figura 74):

> **Professora:** [diz para o Luís, enquanto ele exibe o diagrama de dispersão já com a reta de regressão incluída] Consegues voltar para trás e mostrar os valores de a, b da reta de regressão?
> **Luís:** Sim.
> **Professora:** … Cá está! Tem ali [na tela] os valores a e b e o do coeficiente de correlação que é o *r* da reta que se aproxima de todos os pontos! Que, neste caso, é de 0.99… Isto quer dizer que há uma correlação de que tipo?... Como podemos classificá-la?
> **Vários alunos:** É forte.
> **Professora:** É forte, não é? É quase perfeita! As primeiras alíneas já estão… Agora [nas alíneas seguintes] queremos fazer previsões. Quero utilizar a reta de regressão para, dada uma idade, ver qual a possível altura e vice-versa… (…) Qual a altura do Pedro quando ele tinha 10 anos?
> **Luís:** Não sei [determinar na calculadora]. (observação 8)

Figura 74: Valores obtidos para a expressão analítica da reta de regressão.

Verifica-se que Lia pede aos alunos para caracterizarem o resultado projetado de 0.996 de coeficiente de correlação. Os alunos classificaram esse valor quanto à sua intensidade (“Forte”), não fazendo referência ao seu sinal e também não descreveram a forma como os dados estavam distribuídos no diagrama de dispersão. Com base no valor do coeficiente de correlação, Lia acrescentou apenas que a correlação era “quase perfeita”. Foi também destacada pela professora a utilidade da regressão linear para efetuar previsões, mas não se observou qualquer comentário que explicasse porque é que se poderia usar este modelo.

Lia tinha a expetativa de encontrar, juntamente com os alunos, uma forma rápida de estimar o valor de uma variável sendo conhecido o valor da outra, tendo por base a reta de regressão determinada, tal como as últimas alíneas da tarefa aludiam. Contudo, ao não conseguir descobrir as instruções que pudessem conduzir aos resultados pedidos através das duas calculadoras, acabou por sugerir aos alunos que passassem para o caderno a equação da reta (com *x* variável independente a representar a idade e *y* variável independente a representar a altura) e que os determinassem com papel e lápis. No entanto, os alunos, de uma maneira geral, revelaram mais dificuldades em determinar com papel e lápis o valor de *x* sendo fornecido o valor de *y*, do que o valor de *y* conhecido o valor *x*, tendo por base a mesma equação da reta de regressão. Ao aperceber-se dessas dificuldades, quando circulava de mesa em mesa a ver a resolução dos alunos, optou por ir ao quadro resolver ela própria as duas equações. Assim, a mesma reta de regressão foi usada pela professora e alunos para estimar *y* dado um valor de *x* (10 anos) e *x* dado um valor de *y* (150 cm).

Quanto ao uso das calculadoras gráficas, Lia não estava à espera que não se conseguisse fazer de forma imediata através das calculadoras as estimativas através da reta de regressão. Além disso, mesmo o enunciado da tarefa parece aludir a que se faça uso de uma única reta de regressão. Esta situação sugere que a professora desconhece condições em que as retas de regressão podem ser usadas na previsão de eventos.

Tarefa *Distribuição do número de médicos e farmácias por regiões* (t28)

Esta tarefa foi também escolhida do manual. Em termos de enunciado era semelhante à anterior e a atividade desenvolvida em torno dela foi também análoga. Lia mostrou como encontrar a equação da reta de regressão (com *x* variável independente a ser o número de farmácias e *y* variável dependente a ser o número de médicos) através da calculadora *Texas Instruments* (que tinha sido a última das duas a ser projetada na aula), bem como o valor do coeficiente de correlação. A diferença principal, relativamente à tarefa anterior, emergiu quando a professora sugeriu uma estratégia adicional à alternativa de se fazer à mão a estimativa pedida (do número de farmácias sabendo que o número de médicos era 1464). Então, com a calculadora gráfica, Lia aventou estimar o valor de *x* sendo conhecido o valor de *y* através da interseção de duas retas, a de regressão previamente determinada (y=15.466x-442.88, com r=0.96) e a reta horizontal (y=1464):

> Há uma forma de fazer graficamente… mas para isso não podemos usar o menu *STAT PLOT*, temos de escrever as funções no *menu* habitual [*2ND*+*Y*=]. Agora vou ao *Y*= … É lá que vou colocar a equação da reta de regressão… [diz ao aluno voluntário, Luís, para registar na calculadora, que se encontrava exibida na tela para a turma, a expressão da reta] coloca y=15.466x-442.88… e no outro y abaixo… y= 1464…. agora vamos ver a janela [botão window] para arranjar a escala… (acabando por sugerir mudanças para a maioria dos parâmetros devido ao facto de os valores registados não dizerem respeito às retas introduzidas)… podem sempre experimentar outros valores… vamos fazer a intersecção… (pede ao Luís para mostrar as duas representações gráficas em simultâneo no ecrã da calculadora gráfica e depois pede-lhe para clicar na opção *intersect*, opção 5, do *menu CALC* (através da instrução *2nd*+*CALC*)… na Casio a interseção é imediata! [referindo-se a não ser necessário nessa calculadora gráfica colocar o cursor sobre cada uma das curvas antes de se obter o resultado da interseção]… x=123.29… (observação 8)

Lia indicou as instruções aos alunos para que pudessem fazer a estimativa pedida no enunciado da tarefa através da interseção entre a reta de regressão que se tinha determinado na calculadora gráfica e a reta vertical introduzida de novo. Estas duas retas foram colocadas na lista de funções da calculadora gráfica pelo aluno Luís que estava a trabalhar na calculadora que exibida para a turma. Lia avisou que se deveria ajustar a janela antes de se proceder à interseção desejada, talvez por antecipar dificuldades na obtenção da resposta. Se não o fizesse a interseção teria dado erro e teria de lhes pedir para voltarem à janela para fazerem as alterações nos intervalos de variação de cada parâmetro e só depois encontrarem o valor da interseção. Não foi efetuado nenhum comentário à estimativa obtida nem em torno da limitação deste modelo linear na previsão de eventos.

Salienta-se, ainda, que, nesta tarefa, Lia usou a reta de regressão y em função de x para estimar um valor da variável independente. No entanto, os dados desta tarefa não permitem concluir que as retas de regressão y em função de x e x em função de y coincidam[28], visto o valor absoluto do coeficiente de correlação não ser 1. Assim, não é possível usar a reta de regressão y em função de x para se efetuar previsões para x quando conhecemos um determinado valor para y. A professora revelou mais uma vez desconhecer esta situação, dado que não a colocou em causa. O enunciado da tarefa parece induzir a que se faça precisamente esse erro, por ter nomeado como variável independente o número de farmácias (y) e variável dependente o número de médicos

[28] Geralmente as retas de regressão y como função de x e x como função de y não coincidem, são distintas uma da outra, exceto quando o coeficiente de correlação é 1 ou -1 (Pestana & Velosa, 2009).

(x). A estimação correta deveria ser aproximadamente 149 farmácias em vez das 123. À semelhança da exploração da tarefa anterior, a professora não fez emergir a relação entre o coeficiente de correlação obtido (0.96) e o modelo de regressão linear. Deste modo, não surgiu nenhuma explicação do motivo pelo qual se poderia usar o modelo linear.

6.3.5.3 Reflexões sobre o tópico na aula

Sobre dados bivariados e regressão linear, Lia propôs tarefas que iam ao encontro das capacidades exigidas no programa, achando apenas que, se tivesse mais tempo, teria recorrido a um tipo de atividade, que não realizou, que conduzisse os alunos a trabalhar "mais autonomamente na recolha de dados e análises estatísticas" sobre temáticas que fossem ao encontro dos seus interesses, para poder assumir, como desejava nessas circunstâncias, o "papel de orientadora ou moderadora" (entrevista 2).

O tópico dos dados bivariados e regressão foi o último a ser lecionado na turma de Lia e, como a professora afirmou, acabou por ser daqueles em que teve menos tempo para propor tarefas de consolidação e de prática, nomeadamente, com a calculadora gráfica. Como forma de ganhar tempo, Lia decidiu introduzir todos os conceitos deste tópico essencialmente através da *escola virtual*, como se evidenciou atrás, que incorporava também a visualização de cálculos através da utilização da calculadora gráfica.

Ao observar alguns excertos das suas aulas deste tema, Lia mostrou-se mais focada na sua atitude na interação com os alunos, fazendo alguns comentários: "Estou sempre a fazer perguntas… para a aula não morrer… deveria ser mais calma e dar-lhes mais hipóteses de responderem…" (entrevista 3).

Relativamente aos tópicos da Estatística em causa, Lia lembrava-se que não tinha relacionado o centro de gravidade com a reta de regressão na aula, considerando que esta situação acabaria por ficar esquecida, essencialmente, ao pouco tempo que os nossos próprios pais nunca nos vão deixar sozinhos. Também referiu que, caso colocasse essa questão aos alunos e lhes pedisse para tentarem relacionar os dois conceitos, estes acabariam por conseguir dizer que o ponto centro de gravidade de um conjunto de dados pertenceria à reta de regressão. Na sua perspetiva, abordar a reta de regressão era prioritário bem como utilizar a calculadora gráfica para obter a sua expressão analítica e o seu coeficiente de correlação linear. Como refere, o coeficiente de correlação linear é uma "medida que mede a força da relação [bivariada]… dá para ver se está em concordância com o diagrama de dispersão" (entrevista 2). Assim, acrescenta que um "valor fraco pode indiciar que não haja relação linear mas ela possa ser de outro tipo… mas isto não é trabalhado na Matemática A" (entrevista 2). Lia é da opinião que os alunos ficaram com o conhecimento "de que a reta de regressão é a que se aproxima mais de todos os pontos do diagrama de dispersão" (entrevista 2) e que esta serve também para fazer previsões depois de se obter a sua equação na calculadora gráfica.

6.4.5.4 Discussão do tópico

CETE. Lia apostou na *escola virtual* com vista a uma exposição relativamente rápida de uma seleção de situações que envolviam os conceitos essenciais, o que incluía também mostrar como obter o diagrama de dispersão, o coeficiente de correlação e a reta de regressão na calculadora gráfica. Esses momentos expositivos englobaram alguma interação com os alunos; contudo, houve essencialmente, da parte da professora, preocupação em tecer comentários que visaram o estabelecimento de algumas interrelações entre conceitos e representações apresentados. Na verificação das resoluções dos alunos às tarefas propostas na aula, em grande grupo, a professora teve uma interação limitada com os alunos, não deixando, no entanto, de validar as suas respostas.

A correlação entre duas variáveis foi analisada originalmente pela professora através dos dados representados num diagrama de dispersão e foi descrita pela tendência global observada. A medida usada para quantificar essa correlação foi o coeficiente de correlação, cujo valor foi classificado quanto ao sinal (positivo ou negativo) e intensidade da correlação (tanto maior quanto mais próxima de -1 ou 1). Esta medida foi intuitivamente relacionada com a reta de regressão, nalguns momentos expositivos. Contudo, na resolução das tarefas, esta ideia foi perdendo notoriedade, ou seja, não surgiu nas tarefas de associação entre diagramas e valores de coeficientes de correlação e também não surgiu nas duas tarefas em que os alunos teriam de usar o modelo de regressão linear para fazerem algumas previsões. A atenção dada ao centro de gravidade serviu essencialmente para sustentar uma leitura global dos dados representados no diagrama de dispersão e verificar o sinal da correlação. A interligação entre dados e contexto na interpretação da reta de regressão, dos resultados estimados, ou para realçar limitações do modelo não foi evidente na prática de Lia. O ponderar a fiabilidade do modelo linear ou se a relação entre duas variáveis poderia ser outra que não a linear, nomeadamente, pela observação dos dados representados no diagrama de dispersão, não foi visível no ensino deste tópico.

CA. Lia detetou algumas dificuldades inesperadas nos cálculos que os alunos tiveram de fazer com papel e lápis para determinarem os valores das previsões com o modelo linear. Lia admitiu também ter tido pouco acesso ao raciocínio dos alunos numa ou noutra situação, reconhecendo que os questionou excessivamente sem dar tempo para que pudessem pensar na resposta e verbalizar alguma explicação. Lia considerou que teve pouco tempo para lecionar este tópico e que, consequentemente, não foram dadas mais oportunidades aos alunos para trabalharem mais tarefas e aplicarem a calculadora gráfica.

CE. Lia revelou alguns indícios de possuir um CE mais abrangente do que aquele que evidenciou na sala de aula, nomeadamente quanto a conhecer que o centro de gravidade pertence à reta de regressão e também a relação entre o coeficiente de correlação e o modelo de regressão linear. Lia parece saber que este modelo é obtido através do método dos mínimos quadrados, mas não teceu nenhuma ideia sobre a composição desta reta na aula. A professora revelou desconhecer em que moldes a regressão linear deveria ser efetivamente usada na previsão de eventos. Na aula, Lia propôs uma tarefa que apresentava uma incorreção ao solicitar que se estimasse x, para um dado valor de y, através da reta de regressão y em função de x, quando esta não coincidia com a reta de regressão x em função de y, para assim poder ser usada corretamente. Deste modo, os alunos podem ter ficado com a ideia errónea que qualquer reta de regressão serve não só para estimar valores da sua variável dependente, como também da sua variável independente.

A professora revelou saber como usar os dois modelos de calculadoras gráficas para determinar a reta de regressão e o coeficiente de correlação e ainda para ajustar o ecrã da calculadora (através do menu *window* ou *zoom*) de modo a poder visualizar a representação como desejava, por exemplo. Contudo, mostrou-se surpreendida por não estar a conseguir usá-la para obter diretamente os resultados das previsões solicitadas nas tarefas com o menu *Estatística*.

CC. Neste tópico, Lia fez uma leitura do programa focada nos conceitos e nas representações a lecionar e na utilização da calculadora gráfica. A sua prática revela estar em consonância com essa leitura do currículo. Ao longo das aulas foi possível observar algumas situações em que a opção de não desenvolver o conceito ou de não o relacionar com outros acaba por espelhar também a sua visão do lugar da Estatística no currículo.

6.3.6 Discussão do tema à luz de cada domínio do conhecimento didático

Conhecimento do ensino do tema da Estatística (CETE). Lia revela que a atividade principal das aulas de Estatística, na disciplina de Matemática A do ensino secundário, é a resolução de tarefas pelos alunos. Essa atividade foi bastante orientada pela professora, foi intercalada com vários momentos expositivos realizados pela professora, em que apresentou definições, representações, exemplos de aplicação das noções e de construção das representações, e estabeleceu algumas conexões entre os assuntos referidos. Em todos esses momentos, Lia solicitou a participação dos alunos. Na aula, a calculadora gráfica foi sobretudo usada a partir do momento em que o desvio-padrão foi abordado, ou seja, nos dois últimos tópicos do tema. De todos os tópicos do tema, o tópico *Organização e interpretação de dados* foi aquele em que Lia propôs mais tarefas para serem resolvidas na aula. Nos tópicos *Medidas de localização* e *Medidas de dispersão* a professora não conseguiu sugerir mais tarefas do que aquelas que foram trabalhadas na aula, porque tinha a intenção de promover uma maior participação dos alunos nas análises comparativas de dados propostas, o que a levou a optar por disponibilizar mais tempo letivo para os alunos trabalharem essas tarefas. Nestes tópicos, foi particularmente visível a atenção que a professora deu à variabilidade dos dados, relacionando medidas de localização e de dispersão na comparação de dois conjuntos, e ao estabelecimento de ligações entre Estatística e contexto. Contudo, nessas análises não foram contempladas representações variadas, apenas as mencionadas nas tarefas. Lia fez alguma referência à resistência das medidas de localização, média, moda e mediana, face à eventual presença de algum dado mais afastado na distribuição, não fazendo o mesmo para as medidas de dispersão nem para o coeficiente de correlação.

A professora envolveu os alunos na resolução e correção das tarefas, questionando-os sobre os assuntos. Como, na maioria das vezes, obtinha respostas muito sucintas, acabava por ser ela a explicar o porquê da resposta e fazer comentários às situações que gostaria de ter obtido da parte dos alunos.

Conhecimento do aluno (CA). A professora foi detetando algumas dificuldades dos alunos, na determinação de algumas medidas estatísticas, em construir uma representação gráfica, na interpretação de tabelas de frequências com dados agrupados. Para Lia, todas essas dificuldades eram expetáveis, por serem dificuldades reconhecidas dos outros anos em que lecionou o tema. Já para a vertente interpretativa que tentou que os alunos desenvolvessem na análise dos dados e quando consideravam os resultados estatísticos (nomeadamente, fazer comentários relevantes na comparação de gráficos que diziam respeito a distribuições diferentes, interpretar o valor do desvio-padrão, fazer inferências acerca dos dados a partir dos resultados determinados), Lia não tinha dúvidas que essas atividades eram muito desafiantes para os alunos, apesar de ser a primeira vez que as solicitou em Matemática A. Nestas situações, a professora teve necessidade de completar constantemente as respostas dos alunos, com correções e acrescentando informação.

Conhecimento de Estatística (CE). Lia revela saber usar a calculadora gráfica na obtenção de medidas e algumas representações gráficas. Na aula, alguns conceitos estatísticos foram associados a algumas ideias desadequadas (sondagens e sua credibilidade, média e o seu entendimento como medida representativa, reta de regressão e o entendimento de como esta deve ser usada para estimar valores para as variáveis em estudo). Assim, Lia parece apresentar algumas lacunas ao nível do seu conhecimento estatístico e também alguma falta de experiência em pensar sobre esses assuntos.

Conhecimento do currículo (CC). Para Lia, os assuntos do programa escolar inserem-se na área da Estatística descritiva e enfatizam um desenvolvimento de toda a unidade em torno dos conceitos, cálculos e da construção de representações. A interpretação dos resultados e dos dados, e mesmo a realização de inferências em torno dos resultados e dos dados, não são aspetos explicitamente contemplados no programa. Em termos globais, essas suas ideias sobre a unidade da Estatística no programa estão presentes no trabalho desenvolvido nas suas aulas; no entanto, Lia procurou promover alguma interpretação sobretudo nos tópicos das medidas de localização e de dispersão.

CAPÍTULO 7

CONCLUSÃO

Este capítulo concludente está organizado em três secções. Na primeira apresenta-se uma breve síntese do estudo, focada no seu objetivo principal, nas questões de investigação e na metodologia adotada. Na segunda apresentam-se as conclusões do estudo, estruturadas de acordo com as questões de investigação que o orientaram. Na terceira e última secção, inclui-se uma reflexão final a partir do estudo realizado.

7.1 Síntese do estudo

Este estudo inscreve-se na área de investigação sobre o conhecimento profissional do professor, focando-se no conhecimento didático (Canavarro, 2003; Ponte, 2012; Ponte & Oliveira, 2002) e identificando vários aspetos deste conhecimento fundamentalmente a partir da prática letiva. Assim, o objetivo deste estudo é compreender o conhecimento didático em Estatística de duas professoras na disciplina de Matemática A, no 10.º ano, a partir da análise das suas práticas. Mais concretamente procura-se responder às seguintes questões:

1. Como se caracteriza o conhecimento didático em Estatística das professoras no ensino dos seguintes tópicos do programa:

    - população e amostra;
    - organização e interpretação de dados;
    - medidas de localização;
    - medidas de dispersão;
    - dados bivariados e regressão linear?

2. Que aspetos transversais do conhecimento didático em Estatística das professoras se destacam na articulação entre os tópicos do programa?

O quadro teórico do estudo desenvolve-se em torno do conhecimento profissional do professor de Matemática, nomeadamente do conhecimento didático que diz respeito a um conhecimento orientado para as situações da prática, integrando o pensamento do professor (Canavarro, 2003; Ponte & Oliveira, 2002). Contempla resultados de estudos teóricos e empíricos sobre conhecimento didático do professor em Estatística (por exemplo, Burgess, 2009; Groth, 2013; Oliveira & Henriques, 2014) e perspetivas sobre o ensino e aprendizagem da Estatística (Ben-Zvi & Garfield, 2004; Gattuso & Ottaviani, 2011; Ponte & Fonseca, 2001; Rossman et al., 2006; Scheaffer, 2006) e inclui, ainda, aspetos que dizem especificamente respeito ao ensino e aprendizagem da Estatística, em particular, dos conceitos estatísticos relevantes no contexto do ensino secundário nacional e internacional, tais como variabilidade, população e amostra, medidas de localização e de dispersão e dados bivariados e regressão linear (Batanero et al., 2013; Burrill & Biehler, 2011; Franklin et al., 2007; Garfield & Ben-Zvi, 2008; ME, 2001; Shaughnessy, 2007). O modelo do conhecimento didático em Estatística assumido no estudo envolve quatro domínios que se interrelacionam, nomeadamente conhecimento do currículo, conhecimento do ensino do tema/tópico estatístico, conhecimento do aluno e conhecimento de Estatística.

Esta investigação, de natureza qualitativa (Bogdan & Biklen, 1994; Patton, 2002), segue o paradigma interpretativo (Denzin & Lincoln, 1998; Erickson, 1986), sendo a modalidade escolhida, o estudo de caso (Ponte, 2006; Stake, 2007). Foram elaborados dois casos de duas professoras de Matemática do ensino secundário, com recolha de dados realizada num ambiente naturalista, através de entrevistas, observação de aulas, recolha documental e notas de campo (Bogdan & Biklen, 1994; Savenye & Robinson, 2005). Atendendo à natureza do estudo, a análise de dados foi fundamentalmente descritiva e interpretativa (Bogdan & Biklen, 1994; Huberman & Miles, 1994). Os aspetos que se destacaram em cada uma das professoras relativos aos quatro domínios do conhecimento didático em Estatística não tiveram como objetivo induzir generalizações mas problematizar e trazer um maior entendimento do fenómeno em estudo.

7.2 Conclusões do estudo

7.2.1. O conhecimento didático em Estatística no ensino dos vários tópicos

A primeira questão de investigação, *Como se caracteriza o conhecimento didático em Estatística das professoras no ensino dos seguintes tópicos do programa: População e amostra; Organização e interpretação de dados; Medidas de localização; Medidas de dispersão; Dados bivariados e regressão linear?* é, na verdade, uma multiquestão. Na resposta a esta questão do estudo assume-se a ordem pela qual os tópicos estatísticos nela surgem assinalados, caracterizando-se, para cada um deles, as quatro dimensões do modelo do conhecimento didático em Estatística assumido no estudo, ou seja, conhecimento do currículo (*CC),* conhecimento do ensino do tópico/tema estatístico (*CETE*), conhecimento do aluno (*CA*) e conhecimento de Estatística (*CE).*

7.2.1.1. O conhecimento didático em Estatística no tópico População e amostra

CC. Com este estudo foi possível identificar alguns aspetos do conhecimento do currículo das professoras que dizem respeito à sua leitura dos vários tópicos do programa do ensino secundário vigente no momento (ME, 2001). No que diz respeito ao tópico População e amostra, os dados sugerem que, para as professoras, este tópico inclui assuntos relativamente familiares para os alunos (amostra/população, sondagem/censo) por já terem tido contacto com eles, quer durante o 3.º ciclo, quer na sua vida quotidiana através dos meios de comunicação social. Ambas as professoras encaram este tópico como introdutório ao tema da Estatística do ensino secundário, considerando que estes assuntos devem ser tratados de um modo informal, sem grandes detalhes nem aprofundamento. Lia destaca ainda este tópico como oportuno para se recorrer a dados estatísticos reais, nomeadamente gráficos publicitados nos meios de comunicação social.

CETE. O estudo permite apontar um conjunto de aspetos do conhecimento do ensino do tópico estatístico que as professoras evidenciam em diversos momentos, nomeadamente: quando atendem à sua preparação, através da sequenciação dos conteúdos e escolha das tarefas; e quando conduzem o ensino promovendo, em particular, a compreensão dos conceitos e dos dados e o estabelecimento de conexões envolvendo conceitos abordados, por meio de consideração de *dados e contexto* e utilização de ideias adequadas acerca de amostras.

Todos os assuntos deste tópico foram abordados pelas duas professoras de uma forma bastante sucinta e pouco aprofundada. Na sequenciação de conteúdos, Lia deu algum destaque à noção de amostra e a alguns métodos de amostragem. Estela, por seu turno, não realçou nenhum assunto em particular.

Ambas propuseram poucas tarefas para serem exploradas nas aulas, contudo, os enunciados dessas tarefas mostravam alguns indícios de *dados e contexto*, que é um dos aspetos do raciocínio estatístico que Wild e Pfannkuch (1999) destacam como sendo importante desenvolver em aula. Apesar de Lia ter recorrido a momentos de exposição para apresentar conceitos e situações que envolviam a amostragem e de ter proposto mais tarefas do que Estela, de uma maneira geral, nas suas aulas, não fez emergir questões nem estabeleceu ligações relevantes quanto às características de amostras válidas. Algumas das suas afirmações em aula relativamente à credibilidade de uma amostra nem sempre foram bem conseguidas. Isto pode ter acontecido devido à falta de experiência da professora em raciocinar sobre amostras e em utilizá-las na realização de inferências. Nas tarefas que incluíam representações gráficas houve sobretudo um enfoque na obtenção do resultado através de cálculos básicos ou na realização de reconhecimentos específicos de baixo nível cognitivo de *ler os dados* (Curcio & Artz, 1996; Friel et al., 2001). Foi também evidenciada alguma integração da *Estatística e contexto*, um outro aspeto do raciocínio estatístico destacado por Wild e Pfannkuch (1999). No caso de Estela, de uma maneira geral, a sua intervenção neste tópico foi bastante limitada, bem como a dos seus alunos.

Os alunos do ensino secundário devem ter algum contacto com o processo de recolha de dados para irem reconhecendo e lidando com fontes de *variabilidade nos dados*, com a variabilidade da amostra e com a sua variabilidade não controlada, quer na sua medição, quer na sua observação, com vista ao desenvolvimento do raciocínio estatístico (Franklin et al., 2007). A literatura aponta aspetos basilares das noções estatísticas deste tópico que deveriam ser trabalhados nas aulas. Por exemplo, uma amostra que se pretenda representativa deverá ter propriedades que reflitam as caraterísticas da população; ao utilizar-se uma amostra para se estimar uma caraterística numérica da população, a sua dimensão depende muito da variabilidade da população subjacente (Garfield & Ben-Zvi, 2008; Rubin et al., 1990; Sharma, 2005). A observação de uma maior ou menor variabilidade nos dados tem influência na escolha de uma amostra de maior ou menor tamanho, respetivamente. A opção de se escolher uma amostra de maior dimensão tem de ser ponderada tendo em conta vários aspetos tais como custos, tempo e nível de precisão desejados (Franklin et al., 2007). Os resultados deste estudo sugerem que a prática das professoras no ensino deste tópico é fortemente influenciada pela leitura que fazem do programa (ME, 2001). Além disso, manifestam também algumas das dificuldades associadas ao ensino do tópico e que dizem respeito a argumentar sobre as noções abordadas e a estabelecer algumas relações entre elas.

CA. Relativamente a esta dimensão do conhecimento didático das professoras, o estudo permite tirar conclusões sobre: (i) as suas expectativas em relação à aprendizagem; e (ii) os seus entendimentos acerca do conhecimento desenvolvido pelos alunos nos vários tópicos, face a algumas estratégias adotadas e intervenções na aula. Assim, no presente tópico, as professoras consideraram que os alunos, de uma maneira geral, não teriam dificuldades de compreensão no tópico até porque não tinham o propósito de o aprofundar. Assim, saber distinguir entre população e amostra e ter a ideia de que uma amostra válida teria de ser aleatória e representativa eram os assuntos principais que as professoras consideravam que os alunos deveriam reter do tópico.

Ambas as professoras consideravam que os alunos ponderavam a representatividade de uma amostra através do seu tamanho; contudo, esta estratégia não surgiu de forma convincente nas suas aulas. Por exemplo, Estela apenas nomeou esta estratégia quando refletiu, conjuntamente com a investigadora, sobre o modo como os seus alunos habitualmente entendiam a representatividade de uma amostra.

No que diz respeito a este domínio do conhecimento didático, foram poucas as ocasiões em que as professoras procuraram aceder aos raciocínios dos alunos ou ajudá-los a aperfeiçoar algumas ideias que surgiram da interação nas aulas. Em particular, Lia quando obtinha respostas sucintas dos alunos, geralmente, não lhes pedia explicações. Também Estela não aproveitou as dúvidas de uma aluna quanto à representatividade de uma amostra. Sharma (2005) sugere que a estratégia de pedir explicações aos alunos nas situações que envolvam a representatividade de uma amostra é particularmente útil dado que, muitas vezes, as suas respostas embora corretas apoiam-se em argumentos desadequados. Também nem sempre foram detetadas por Lia algumas imprecisões nas respostas dadas pelos alunos. Esta situação pode estar relacionada com a necessidade de algum aprofundamento ao nível do seu conhecimento estatístico. Os dados também sugerem que o CA das professoras é bastante influenciado pela sua leitura do programa da disciplina.

CE. O estudo permite destacar alguns aspetos deste domínio quanto à compreensão das noções e ao estabelecimento (ou não) de relações entre as noções abordadas nas aulas (*raciocínio com modelos*). As duas professoras revelaram conhecer alguns métodos de amostragem e ter o conhecimento de que quanto maior for o tamanho das amostras mais probabilidade têm de ser representativas, o que constituí um exemplo de *raciocínio com modelos* (Burgess, 2009; Wild & Phankuch, 1999), um dos aspetos do raciocínio estatístico. Contudo, Lia insistiu na preferência por recenseamentos em oposição às sondagens, devido ao tamanho da amostra e fiabilidade de resultados, parecendo não conhecer situações em que o recurso à amostragem aleatória dá resultados fidedignos e com margens de erro muito pequenas. Estela, por seu turno, considerava correto que se averiguasse a possibilidade de uma amostra ser representativa dando sobretudo relevo à sua dimensão. No contexto escolar, essas situações podem fazer com que os alunos desenvolvam algumas ideias erróneas acerca das amostras representativas, o que está em sintonia com os resultados de Rubin et al. (1990) e Sharma (2005) que apontam a necessidade dos alunos do ensino secundário perceberem as caraterísticas das boas amostras e o que pode constituir uma amostra enviesada, e mais geralmente, adquirirem noções adequadas sobre representatividade e variabilidade de uma amostra.

7.2.1.2. O conhecimento didático em Estatística no tópico Organização e interpretação de dados

CC. Os dados sugerem que as professoras, tendo em conta a sua leitura do programa do ensino secundário, consideram que este tópico é claramente centrado na obtenção das representações tabulares (que incluem frequências) e gráficas e nalgumas leituras, embora básicas, em torno dessas representações.

CETE. O estudo permite identificar um conjunto de aspetos do conhecimento do ensino deste tópico estatístico que as professoras evidenciam em diversos momentos, relativamente: à preparação do tópico; e à condução do ensino promovendo, em particular, a compreensão dos conceitos e dos dados e o estabelecimento de conexões envolvendo conceitos e representações (utilização da calculadora gráfica; uso da *transnumeração*; *integração da Estatística e contexto*).

No ensino deste tópico, as duas professoras mencionaram os mesmos conteúdos, destacando as percentagens e as frequências absolutas/relativas acumuladas, e realizaram uma sequenciação semelhante dos assuntos tratados, embora o tenham explorado de uma forma diferenciada. Lia insistiu mais na construção de tabelas de frequências, a partir de dados representados em representações tabulares e gráficas diversificadas, propondo várias tarefas com esse objetivo principal, mas não solicitou a utilização da calculadora gráfica. Estela, por seu turno, propôs várias tarefas que tinham como propósito alcançar representações tabulares ou gráficas variadas, solicitando que a calculadora gráfica fosse, sempre que possível, usada para obter essas representações. O incentivo dado por Estela ao uso da calculadora gráfica tinha a ver com a sua perspetiva de que esta é uma ferramenta útil para a determinação das representações e que facilita o trabalho autónomo dos alunos. Na opinião de Lia, a utilização da calculadora gráfica só faria sentido se tivesse mais tempo letivo para dedicar a este tópico, considerando que uma aprendizagem eficiente passava por colocar os alunos a realizar os cálculos e tabelas de frequências manualmente.

Nas aulas das duas professoras houve acesso a dados (sobretudo quantitativos), a representações gráficas variadas (tais como gráfico circular, barras, histograma, pictogramas, diagramas de caule e folhas, função cumulativa) e a representações tabulares (com frequências). No entanto, o questionamento acerca dos dados e do seu contexto nessas representações e o reconhecimento de caraterísticas específicas das representações estiveram pouco presentes. Assim, a *transnumeração* realizada na aula não teve o propósito de se chegar a um melhor entendimento dos dados e da sua variabilidade, e dessa forma contribuir para o raciocínio estatístico dos alunos (Burgess, 2009; Shaughnessy, 2007). Esta constituiu, sobretudo, uma forma de os alunos praticarem a construção de representações, no caso de Lia, e de determinação das mesmas através da calculadora gráfica, no caso de Estela. Ainda, no caso de Lia, foi possível observar um diálogo interativo, embora curto, entre a professora e alunos sobre uma tabela de frequências que mostrava o potencial de os questionar sobre os dados e seu contexto no sentido de os orientar para análises mais ricas. Também a conexão estabelecida entre frequências acumuladas e função cumulativa teve um propósito, essencialmente, procedimental nas aulas das duas professoras. A função cumulativa foi exemplificada com uma situação concreta que exibia uma função cumulativa (analítica e graficamente) e consolidada através de uma tarefa que serviu para mostrar como converter uma tabela de frequências absolutas acumuladas numa função matemática, nas versões expressão analítica ou representação gráfica. Acima de tudo, neste tópico foi proporcionado aos alunos um contacto com uma variedade de representações. Curcio e Artz (1996), Friel et al. (2001) e Garfield e Ben-Zvi (2008) salientam que o enriquecimento das tarefas com questões que conduzam à comparação de representações pode facilitar o desenvolvimento da compreensão dos conceitos envolvidos, o reconhecimento da utilidade das representações e o entendimento dos dados e contexto. Essas atividades, além de permitirem o desenvolvimento de capacidades nos alunos de nível de intermédio/avançado no campo da Estatística, têm potencial para promover o raciocínio estatístico (Franklin et al., 2007; Garfield & Ben-Zvi, 2008).

CA. No que diz respeito às expectativas em relação à aprendizagem, as professoras consideravam que os alunos, de uma maneira geral, não teriam grandes dificuldades em trabalhar os assuntos deste tópico, dado que o que lhes iria ser proposto tinha a ver com a resposta a situações com questões estruturadas e focadas na construção de representações que envolviam frequências. No entanto, ao nível do conhecimento desenvolvido pelos alunos, foram detetando algumas dificuldades dos alunos em: *ler os dados* através de tabelas e gráficos; relacionar a informação das tabelas com o contexto, ou seja, em *ler entre dados*; e usar essa informação para responderem a algumas questões incorporadas nas tarefas, isto é, em *ler além dos dados* (Curcio et al., 1996). Por exemplo, Lia revelou-se surpreendida pelas dificuldades que a maioria dos alunos teve em perceber que o gráfico de barras, fornecido numa tarefa, representava uma distribuição com as frequências absolutas acumuladas. Os alunos além de não estarem a entender o que significava cada barra do gráfico, não sabiam como converter essas frequências acumuladas em frequências simples para poderem responder às questões da tarefa, apesar de terem tido uma experiência prévia com construção de várias tabelas de frequências que envolviam as frequências acumuladas. Lia reconheceu que não bastava colocar os alunos a consolidar a construção de tabelas de frequências para que ficassem capazes de fazer leituras de dados corretas sobre uma qualquer representação gráfica que envolvesse frequências.

Estela também detetou algumas dificuldades na utilização da calculadora gráfica para se chegar às representações pedidas e não se apercebeu de alguns erros cometidos por alunos na representação de frequências relativas na calculadora gráfica.

Ainda assim, as professoras poucas vezes questionaram os alunos acerca dos dados e sobre eventuais conexões entre representações e, por isso, tiveram um acesso algo limitado aos seus raciocínios.

CE. O estudo permite destacar alguns aspetos deste domínio quanto ao estabelecimento (ou não) de relações envolvendo conceitos e representações deste tópico (uso da *transnumeração* e *integração da Estatística e contexto*). No ensino deste tópico, foi possível observar o à vontade de Estela em ajudar os alunos a chegar às representações tabulares e gráficas especialmente através da calculadora gráfica, conhecendo os diferentes parâmetros a terem que ser obrigatoriamente ajustados e sequências de passos para as alcançar e, ainda, desbloqueando qualquer erro com que os alunos se confrontassem na calculadora. Lia, por sua vez, deu particular atenção à construção manual de tabelas de frequências, optando por não usar a calculadora gráfica. Ambas as professoras revelaram conhecer representações tabulares e gráficas variadas mas não privilegiaram análises comparativas entre representações, não fazendo uso da *transnumeração* e da *integração da Estatística e contexto* (Wild & Pfannkuch, 1999). No entanto, estas são algumas das capacidades que o professor deve mobilizar no ensino da Estatística, nomeadamente quando exibe representações diversificadas na aula (Ben-Zvi & Garfield, 2005; Burrill & Biehler, 2011; Franklin et al., 2007).

7.2.1.3. O conhecimento didático em Estatística no tópico Medidas de localização

CC. Este estudo evidencia alguns aspetos do conhecimento curricular que dizem respeito à leitura do programa escolar no tópico. A leitura de Estela centra-se nos aspetos processuais, em que a calculadora gráfica tem um papel fundamental na determinação das medidas de localização e na representação dos diagramas de extremos e quartis. Lia destaca a importância de se saber estas noções e de se construir as representações gráficas manualmente. A professora considera que a calculadora gráfica pode ter um papel importante na confirmação dos resultados e em situações que incorporam amostras com muitos dados.

CETE. O estudo permite identificar um conjunto de aspetos do conhecimento do ensino que as professoras evidenciam em diversos momentos, nomeadamente: quanto à preparação do tópico; e quanto à condução do ensino promovendo, em particular, a compreensão dos conceitos e dos dados e o estabelecimento de relações envolvendo conceitos e representações (utilização da calculadora gráfica; *transnumeração*; consideração da *variabilidade*; *integração da Estatística e contexto*).

Neste tópico, a atividade das duas professoras na aula centrou-se na orientação e correção do trabalho dos alunos em torno de tarefas diversificadas. Realizaram uma sequenciação de conteúdos semelhante embora, no caso de Lia, fosse visível uma maior incidência em dados agregados. Nas aulas de Estela, destacam-se as tarefas que selecionou de diferentes manuais diferentes e, no caso de Lia, evidenciam-se as que foram criadas ou adaptadas por si, bem como algumas que foram selecionadas do manual adotado. Impulsionadas pelas conversas que tiveram com a investigadora e pelas reflexões que realizaram ao longo do estudo, optaram por trabalhar nas suas aulas de Estatística algumas tarefas que nunca tinham proposto aos seus alunos no ensino deste tema. Mais concretamente, no caso de Estela essas tarefas envolviam a comparação de gráficos e, no caso de Lia, a interpretação de resultados das medidas de localização e a análise comparativa de gráficos, com comunicação de descrições e realização de inferências sobre os dados. Estas tarefas tinham potencialidade para desenvolver o raciocínio estatístico dos alunos, sobretudo ao nível da consideração da *variabilidade*, *transnumeração* e *integração da Estatística e contexto* (Burrill & Biehler, 2011; Curcio & Artz, 1996; Garfield & Ben-Zvi, 2005; Garfield & Ben-Zvi, 2008).

Vários autores (por exemplo, Groth & Bergner, 2006; Langrall & Mooney, 2010; Shaughnessy, 2007) sugerem que o ensino das medidas de localização se desenvolva quer ao nível processual quer ao nível conceptual, e que as ideias informais sobre essas medidas partam de expressões mais familiares (por exemplo, *meios*, *a maior parte*) para conceitos mais elaborados (*valor típico, ponto de equilíbrio, sinal no meio do ruído*). Salientam, ainda, que estas medidas devem ser comparadas entre si, confrontadas quando dois ou mais conjuntos de dados são objeto de análise, relacionadas com outras noções e analisadas em representações gráficas, explicando o que transmitem sobre os dados.

Com Estela, as medidas de localização foram abordadas sobretudo ao nível processual. A professora insistiu no uso da calculadora gráfica para obter medidas e representações na resolução de todas as tarefas propostas, embora a interpretação desses resultados não tivesse sido privilegiada, mesmo quando surgiram referências a alguma dessas medidas em representações gráficas. Em particular, não foi visível nas práticas de Estela a interpretação da média como *ponto de equilíbrio* nem a interpretação das medidas de localização, média, moda e mediana, como *sinais no meio do ruído* (Garfield & Ben-Zvi, 2008). Além disso, a *integração da Estatística e contexto*, aspeto do pensamento estatístico destacado por Wild e Pfannkuch (1999) não teve reflexo nas práticas da professora. Quanto aos quartis, do ponto de vista teórico, na aula, surgiu a ideia de que os quartis dividem a distribuição em quatro partes iguais. Contudo, não foi feita referência explícita às estratégias que levam à partição do conjunto de dados, de modo a determinar-se os valores dos quartis, e não foi valorizada a observação e interpretação desses valores à luz dos dados e contexto. Para determinar o valor de um quartil os alunos ficaram então dependentes da calculadora gráfica. Esta situação constitui também um exemplo que mostra uma maior ênfase nos processos no ensino de Estela. A análise comparativa de dois conjuntos de dados não foi central nas suas aulas. A conexão entre mediana e variabilidade, combinando capacidades ao nível *ler entre os dados* e *ler por detrás dos dados* (Shaughnessy, 2007), foi apenas observada na resolução de uma tarefa em que se tentou estabelecer a correspondência entre histograma e diagrama de extremos e quartis de uma mesma distribuição. Assim, não se verificou o uso da *transnumeração*, no sentido atribuído Wild e Pfannkuch (1999).

No caso de Lia, as medidas de localização foram abordadas e desenvolvidas de uma forma que promovia alguma compreensão conceptual, nomeadamente como medida de *centro* e como *sinal no meio do ruído* (Garfield & Ben-Zvi, 2008). Foi dada *atenção à variabilidade* e à *integração da Estatística e contexto* (Wild & Pfannkuch, 1999) na análise dos dados e interpretação dos resultados, em situações de comparação de dois conjuntos de dados. Nas tarefas que desenvolveu para estas aulas, Lia considerou a comparação de conjuntos de dados como uma atividade particularmente útil para desenvolver noções e representações associadas ao tópico e colocar os alunos a fazer comentários sobre os resultados, no sentido de promover análises com *integração da Estatística e contexto* (Wild & Pfannkuch, 1999). Por exemplo, isto aconteceu mesmo quando abordou o diagrama de extremos e quartis, em que solicitou a comparação entre dois diagramas que correspondiam a distribuições de dados diferentes de uma mesma situação problemática, com o intuito de proporcionar entendimento ao nível de *ler entre os dados* e ao nível de *ler além dos dados* (Curcio & Artz, 1996; Shaughnessy, 2007). Na exploração dos dados nessas tarefas, a articulação das ideias de centro e variabilidade foi, muitas vezes, evocada pela professora como um exercício obrigatório para facilitar a compreensão dos dados; contudo, a capacidade de *transnumerar* (Wild & Pfannkuch, 1999) nessas comparações acabou por estar presente mas sem que os alunos pudessem escolher ou elaborar as representações que achassem mais convenientes para as suas análises. Precisamente ao longo dessas situações, Lia forneceu explicações relativamente às medidas de localização que correspondiam ao nível *abstrato estendido* do modelo SOLO descrito por Groth e Bergner (2006). A professora não introduziu a calculadora gráfica neste tópico, ainda que alguns alunos a tivessem utilizado para efetuarem os cálculos básicos.

Os dados agrupados em classes foram essencialmente mencionados nos momentos expositivos conduzidos pelas professoras, nos quais mostraram localizações da mediana e quartis nalgumas situações e relacionaram algumas representações gráficas, embora sempre de uma forma expedita. Este assunto também teve pouca expressão nas tarefas propostas aos alunos. Lia contemplou mais aspetos relacionados com o raciocínio estatístico (Wild & Pfannkuch, 1999), como a atenção à *variabilidade* e *integração da Estatística e contexto*, do que Estela, aspetos esses que emergiram nas suas aulas de uma forma consistente. Identificaram-se algumas situações que espelham também a pouca experiência que as duas professoras têm em combinar ideias entre medidas de localização e variabilidade dos dados, no ensino, em contextos de análises de dados. Resultados semelhantes foram encontrados nos estudos relacionados com medidas de centro de Groth e Bergner (2006) e de Watson e Callingham (2013). Os resultados do presente estudo estão também em sintonia com os de outros estudos que identificaram dificuldades na compreensão das medidas de localização quer nos alunos quer nos professores (Ben-Zvi & Garfield, 2004). Adicionalmente, o facto de as professoras terem decidido explorar algumas tarefas que nunca tinham proposto em aula, parece ter contribuído, embora de modo diferenciado, para alguma ampliação do conhecimento curricular de cada uma delas no ensino do tópico.

CA. As duas professoras consideraram que as noções básicas e representações gráficas associadas ao tópico foram devidamente aprendidas pelos alunos, correspondendo globalmente às suas expectativas. Na aula de Lia acabou por haver uma maior participação dos alunos, talvez devido às questões das tarefas, que apelavam a alguma discussão. Ainda assim, tal como esperava, Lia notou que, na maioria das vezes, os alunos não sabiam por onde começar as suas análises e em que aspetos deviam fazer incidir os seus comentários. Nessas interações que teve com os alunos, Lia completou constantemente as suas respostas, chamando a atenção para as situações que considerava que não deveriam descurar, nomeadamente, a relação entre medidas de localização e variabilidade global dos dados e a análise comparativa da variabilidade tomando diferentes intervalos de variação dos dados.

Lia detetou algumas dificuldades dos alunos na leitura de tabelas de frequências com dados agrupados em classes, o que condicionou à partida a determinação do valor correto de média. A professora considerava que esta dificuldade poderia ser ultrapassada se tivesse mais tempo para lecionar o tópico, propondo mais tarefas que contemplassem leitura e interpretação de resultados. Contudo, também pôde constatar várias dificuldades dos alunos na realização de análises de dados a um *nível de desenvolvimento elementar/intermédio* (Franklin et al., 2007). Nenhuma dessas dificuldades se prendeu com a utilização da calculadora gráfica, mas com a determinação/identificação dos valores corretos dos quartis, construção do diagrama de extremos e quartis e com a realização de descrições adequadas às distribuições de dados, a partir dos valores dos quartis e das medidas de localização, e à luz do contexto dos dados.

Para Estela era essencial que os alunos adquirissem a capacidade de utilizar a calculadora gráfica para alcançar qualquer resultado ou representação gráfica na exploração das situações relacionadas com o tópico. Estela tirou várias dúvidas acerca da utilização do *menu* de Estatística da calculadora gráfica (identificar os dados a colocar nas listas e como introduzi-los; verificar se foram utilizadas as listas corretas nas instruções; proceder ao ajustamento de diversos parâmetros no *zoom* e como fazer para visualizar a representação desejada; confirmar a sequência de instruções, etc.). A interpretação dos resultados das medidas de localização não foi explorada pelos seus alunos e a interpretação do diagrama de extremos e quartis com vista à análise da variabilidade e entendimento dos dados foi realizada numa situação; contudo, de uma forma pouco aprofundada.

De uma maneira geral, quer nas aulas de Estela quer nas de Lia, as respostas corretas dadas pelos alunos foram muito concisas e as professoras não procuraram outras formas de ter acesso aos seus raciocínios. As duas professoras também ficaram surpreendidas com os erros que encontraram na análise dos registos escritos dos alunos referentes a resoluções de algumas tarefas propostas para trabalho de casa. No caso de Estela, vários alunos revelaram dificuldades em chegar aos valores corretos dos quartis e nalgumas situações revelaram não perceber o significado do valor do quartil. Esta situação vem na linha do que referem outros autores (Garfield & Ben-Zvi, 2008; Shaughnessy, 2007), que consideram que, no contexto escolar, há que ir além da esfera computacional e promover um entendimento conceptual que inclua várias ideias sobre cada noção, bem como relações entre elas. Lia, por seu turno, tinha a expetativa de que os alunos teriam um melhor desempenho nas questões que envolviam análises comparativas de dados, visto terem sido trabalhadas algumas situações nas aulas. Contudo, reconheceu dificuldades similares às identificadas nas aulas, mas, ainda assim, não deixou de sublinhar o potencial dessas análises na aprendizagem.

CE. O estudo permite destacar alguns aspetos deste domínio quanto à compreensão dos conceitos e dos dados e quanto às relações estabelecidas envolvendo conceitos e representações (atenção à *variabilidade*). As medidas de localização foram inicialmente descritas pelas professoras de uma forma marcadamente processual. Ambas denotaram a pouca resistência da média face a conjuntos de dados com eventuais *outliers* ou valores extremos, dando preferência à mediana nessas circunstâncias; contudo, esta caraterística da média não emergiu na aula de Estela.

Talvez pela falta de experiência das professoras não foi evidente nas suas práticas o estabelecimento de ligações entre medidas de localização e os conjuntos de dados (por vezes, representados graficamente) com mais descrições e argumentos adequados, em contextos que envolviam alguma atenção à variabilidade dos dados fornecidos. Tal como evidenciado nos estudos de Groth e Bergner (2006) e de Watson e Callingham (2013), revela-se a importância de um conhecimento estatístico que integre as ideias de centro e variabilidade e suas ligações enquanto se realiza análises de dados, para melhor articular essas mesmas ideias no ensino. Isto está também em sintonia com a recomendação que vários autores fazem relativamente ao facto de os professores necessitarem de possuir experiências diversificadas no âmbito da análise de dados (Gattuso & Ottaviani, 2011; Snee, 1993), de modo a desenvolverem gradualmente o seu raciocínio e pensamento estatísticos.

7.2.1.4. O conhecimento didático em Estatística no tópico Medidas de dispersão

CC. Para as professoras, a medida de dispersão mais relevante é o desvio-padrão, cujo valor deve ser determinado na calculadora gráfica. A interpretação deste resultado à luz dos dados e contexto (Wild & Pfannkuch, 1999) não é um requisito do programa, e não é usualmente contemplada nas suas práticas.

CETE. Neste domínio, o estudo permite identificar um conjunto de aspetos que as professoras evidenciam em vários momentos, relativamente: à preparação do tópico; e à condução do ensino promovendo, em particular, a compreensão dos conceitos e dos dados e o estabelecimento de conexões envolvendo conceitos e representações (utilização da calculadora gráfica; consideração da *variabilidade*; utilização da *transnumeração*; *integração da Estatística e contexto*).

Alguns estudos revelam que, dentro das medidas de dispersão (nomeadamente, amplitude total, amplitude interquartil e desvio-padrão), a mais difícil para os alunos será o desvio-padrão, não tanto ao nível processual mas na interpretação do seu valor (Ben-Zvi & Garfield, 2008; delMas & Liu, 2005). A aquisição de uma compreensão significativa do desvio-padrão passa, em particular, por relacioná-lo com os conceitos a partir dos quais este se constrói (nomeadamente, de média e desvio à média) e também por associá-lo à medida da média dos valores absolutos dos desvios. Essa compreensão deve também ser apoiada em leituras dos dados representados graficamente e na reflexão acerca da resistência do desvio-padrão (delMas & Liu, 2005; Garfield & Ben-Zvi, 2008).

Nas aulas de Estela, os valores da amplitude total e da amplitude interquartil não surgiram explicitamente como medidas de dispersão, mas como medidas estatísticas que podiam ser calculadas a partir da observação dos dados ou a partir do diagrama de extremos e quartis. No caso de Lia, foram consideradas como medidas de dispersão que também poderiam suportar leituras realizadas a diagramas de extremos e quartis. Na sequenciação dos conteúdos neste tópico, as duas professoras deram especial destaque ao desvio-padrão e a algumas das suas propriedades. Para as duas professoras, o desvio-padrão foi também assumido como uma medida que não pode deixar de ser referida sem se fazer alusão à média, tal como recomendado por vários autores (Garfield & Ben-Zvi, 2008; delMas & Liu, 2005; Makar & Confrey, 2005). Adicionalmente, no caso de Lia, o desvio-padrão foi considerado a medida de dispersão mais interessante e relevante. A professora apontou algumas das limitações dessas medidas ao nível da informação que podia dar sobre os dados.

A referência à fórmula do desvio-padrão que Estela e Lia fizeram nas aulas pareceu importante para dar ênfase à relação da medida com a média, nomeadamente, através dos desvios à média (incorporados na fórmula). Serviu também para reforçar a ideia de que o desvio-padrão diz respeito a uma medida que reporta a variabilidade dos dados considerados face à média. A ideia intuitiva geral de desvio-padrão apresentada pelas professoras foi que se tratava de uma medida que media a variabilidade ou desvios dos dados relativamente à média. Contudo, enquanto Lia manteve esse mesmo raciocínio ao longo das aulas, Estela tomou o desvio-padrão mais frequentemente como uma medida que fornece uma certa distância entre dado e média, acabando por transmitir uma ideia não totalmente válida. Lia também referiu o desvio-padrão como uma medida útil para indicar diferenças nas distribuições mas não sustentou esse conhecimento recorrendo à *transnumeração* (Wild & Pfannkuch, 1999).

Desafiadas pelas conversas que tiveram com a investigadora e pelas reflexões que realizaram ao longo do estudo, as professoras procuraram incluir tarefas que abarcassem alguma comparação de dados, Lia através de conjuntos de dados numéricos e Estela através de representações diversificadas. Lia usou algumas tarefas com questões equivalentes através das quais foi recorrendo à comparação de dois conjuntos de dados, dando atenção na análise desses conjuntos, à variabilidade, interpretando resultados de medidas estatísticas, ligando *Estatística e contexto*. No entanto, não promoveu a integração de representações gráficas nessas análises (*transnumeração*). Usou pela primeira vez o *menu* Estatística da calculadora gráfica para ensinar a calcular o desvio-padrão e mostrar que, através da calculadora, seria possível conferir todos os outros valores de medidas estatísticas trabalhadas no tema. Na seleção do valor de desvio-padrão obtido na calculadora gráfica, não distinguiu desvio-padrão amostral do populacional, solicitando que se retirasse da calculadora gráfica o que correspondia ao desvio-padrão populacional.

Estela optou por propor aos alunos a resolução de algumas tarefas diversificadas, com recurso à calculadora gráfica. Centrou-se na obtenção dos resultados e das representações, não privilegiando o relacionamento entre eles e a realização de interpretações. De facto, o enfoque incidiu na utilização da calculadora gráfica para chegar-se ao resultado, tal como no estudo de delMas e Liu (2005). Deste modo acabou por condicionar o nível de exigência cognitiva dessas tarefas, que tinham potencial para estimular o raciocínio estatístico dos alunos, em particular, o desenvolvimento conceptual do desvio-padrão (delMas & Liu, 2005; Garfield & Ben-Zvi, 2008). Embora Estela tenha dado uma atenção bastante residual à *variabilidade*, em relação ao desvio-padrão, esse cuidado nem sempre foi explicado da forma mais adequada. Estela fez a distinção dos dois valores de desvio-padrão representados na calculadora gráfica (amostral/populacional) e solicitou que os alunos retirassem da calculadora o resultado de desvio-padrão que estivesse em conformidade com a proveniência dos dados, ou seja, de uma amostra ou população, respetivamente.

Os dados do estudo sugerem que as professoras têm alguma dificuldade em articular de um modo mais profundo noções e representações no ensino deste tópico, o que é também salientado no estudo de delMas e Liu (2005). Ambas demonstram ainda pouca experiência em lidar com medidas de dispersão nas tarefas que trabalharam e parecem desconhecer contextos diversificados para explorar o tópico, à semelhança dos resultados do estudo de González (2013). Adicionalmente, o facto de as professoras terem decidido propor no tópico algumas tarefas que nunca tinham experienciado junto dos alunos parece ter contribuído, embora de modo diferenciado, para alguma ampliação do seu conhecimento curricular.

CA. Para Lia, o maior desafio no ensino deste tópico consistiu em colocar os alunos a comparar resultados e a comentá-los, enquanto Estela considerava que os alunos, de uma maneira geral, não iriam encontrar grandes dificuldades na exploração das tarefas na aula, a partir do momento que percebessem como utilizar a calculadora para obterem os resultados.

Ambas as professoras consideraram que os alunos, em geral, adquiriram o conhecimento de que o desvio-padrão se tratava de uma medida particularmente útil para avaliar a dispersão global dos dados face à média. Ambas estiveram atentas às dificuldades manifestadas pelos alunos ao longo da resolução das tarefas sugeridas na aula, ajudando-os a superar desafios na utilização da calculadora gráfica no cálculo do desvio-padrão. Apesar de solicitar aos alunos o cálculo manual do desvio-padrão, Lia também identificou que tinham dificuldades em aplicar corretamente a fórmula.

As duas professoras acabaram por ter, de uma maneira geral, um acesso limitado ao raciocínio dos alunos neste tópico. Contudo, os alunos de Lia acabaram por fornecer mais algumas explicações sobre os resultados alcançados, devido à natureza das questões propostas nas tarefas e, pelo facto, de a professora insistir em ouvir algumas respostas. Na aula de Estela, de uma maneira geral, não foram levantadas questões nem os alunos foram interpelados sobre os cálculos obtidos. Esta professora também reconheceu que os alunos não desenvolveram a capacidade de interpretar algum resultado ou representação no tópico, apesar de oralmente na aula ter feito alguns comentários deste tipo nalgumas situações isoladas.

CE. O estudo permite destacar alguns aspetos deste domínio quanto à compreensão de dados e de alguns conceitos e quanto às relações estabelecidas envolvendo conceitos e representações (consideração da *variabilidade*). As professoras consideram que o desvio-padrão é a medida de dispersão mais relevante dentro das lecionadas, indicam a sua fórmula, assinalando as parcelas referentes aos desvios à média, e descrevem-no informalmente de uma forma global, embora correta, como uma medida que mede a variabilidade dos dados face à média. Ambas revelam saber usar o *menu* Estatística das calculadoras para determinar o desvio-padrão, e conhecer a pouca resistência desta medida face a *outliers*, mas não referiram essa propriedade na aula. Estela referiu, por vezes, uma outra noção informal de desvio-padrão, como indicador da distância entre dado e média, que não é adequada visto não ser uma regra geral.

7.2.1.5. O conhecimento didático em Estatística no tópico Dados bivariados e regressão linear

CC. Para as professoras, o programa centra-se na análise da relação bivariada através da leitura do diagrama de dispersão, da determinação das coordenadas do centro de gravidade, do cálculo do coeficiente de correlação e da obtenção da reta de regressão (para proceder a eventuais previsões), sem prescindir da realização de cálculos com apoio da calculadora gráfica.

A relação bivariada abordada no tópico diz respeito a uma das seguintes situações: ou não existe, ou, se existe, é traduzida linearmente. Ambas as professoras assumem que qualquer situação que encontrem num manual de Matemática A do 10.º ano (ME, 2001) propondo a análise duma relação bivariada, num contexto de modelação, anuncia obrigatoriamente uma relação linear entre duas variáveis. Os dados do estudo sugerem que as professoras não recorrem à *transnumeração* (Wild & Pfannkuch, 1999) por entenderem a relação bivariada desta forma, centrando-se em análises gráficas dos dados e das suas componentes menos ricas, e não dando destaque a leituras mais pormenorizadas dos dados (leituras por variável, locais e globais dos dados). Esse entendimento da relação bivariada parece também condicionar a possibilidade de os alunos desenvolverem a noção do coeficiente de correlação, ou seja, dando atenção à sua pouca resistência a *outliers* e a fatores que possam influenciar o valor deste coeficiente. A consideração de todos estes aspetos é relevante para melhorar o conhecimento do aluno, no tópico, ao nível processual e conceptual (por exemplo, Anscombe, 1973; Cobb et al., 2003; Engel & Sedlmeier, 2011; Garfield & Ben-Zvi, 2008).

Este tópico é também o último do programa e é, por conseguinte, aquele que tende a ser lecionado de uma forma ainda mais rápida, apesar de ser um tópico novo para a maioria dos alunos. Esta situação que surgiu na prática das duas professoras informa também acerca da visão que as professoras detêm do lugar da Estatística no currículo do ensino secundário e tem repercussões ao nível do CETE de cada professora e da aprendizagem dos alunos.

CETE. O estudo permite identificar um conjunto de aspetos do conhecimento do ensino que as professoras evidenciam em vários momentos, nomeadamente: quando atendem à preparação deste tópico; e quando conduzem o ensino promovendo, em particular, a compreensão dos conceitos e dos dados e o estabelecimento de relações envolvendo conceitos e representações (utilização da calculadora gráfica; consideração da *variabilidade*; utilização do modelo linear e realização de previsões; *integração da Estatística e contexto*).

A sequenciação de conteúdos foi feita, pelas professoras, de uma forma diferenciada. Para Estela e Lia, o diagrama de dispersão era uma representação gráfica fundamental que suportava todo o trabalho a desenvolver no tópico. Estela pretendia que os alunos abordassem as situações propostas a partir da calculadora gráfica. Assim, no início do tópico ensinou como a utilizar para obter representações gráficas e numéricas associadas aos dados bivariados. No seu caso, a reta de regressão surgiu antes da análise dos dados relativamente à associação ou correlação entre duas variáveis. Já Lia favoreceu inicialmente essas mesmas análises antes de explicar como determinar o valor do coeficiente de correlação e os coeficientes da reta de regressão na calculadora gráfica. As duas professoras contemplaram a realização de previsões no final do tópico.

No ensino deste tópico, as professoras optaram por uma abordagem direta desempenhando um papel ativo na condução das atividades e na avaliação do trabalho realizado pelos alunos. Lia recorreu ao sítio virtual *e-escola* para apresentar situações e definições dos conceitos, antes de propor as tarefas que tinha selecionado. Por sua vez, Estela optou por intercalar alguns momentos expositivos no decorrer da resolução das tarefas por si escolhidas para levar à sala de aula. Ambas as professoras escolheram algumas tarefas diversificadas de manuais que incluíam apenas um conjunto de dados bidimensionais e abarcavam questões que, além de semelhantes, eram bastante estruturadas, e ainda outras que incluíam nuvens de pontos/diagramas de dispersão e valores numéricos de coeficientes de correlação para que fosse estabelecida uma correspondência entre gráfico e valor fornecido.

Relativamente aos conteúdos trabalhados nas aulas, no ensino deste tópico, foram sobretudo privilegiadas referências gerais aos conceitos com base nalgumas relações: (i) a possibilidade de associação entre dados bivariados quando no diagrama de dispersão se observa uma tendência global de crescimento ou decrescimento; (ii) a ideia central de reta de regressão como *aquela que melhor se ajusta ao conjunto de dados*; (iii) a ideia do valor do coeficiente de correlação como uma medida que dá informações acerca do sinal (positivo/tendência crescente ou negativo/tendência decrescente) e da força da relação bivariada (quanto mais forte mais próxima do valor -1 ou 1); (iv) a noção de que a reta de regressão pode ser usada como um modelo para fazer previsões; e (v) a ideia de que os valores dos coeficientes da reta de regressão e o valor do coeficiente de correlação devem ser determinados através da calculadora gráfica. O centro de gravidade foi apresentado pelas duas professoras como um ponto de coordenadas $(\bar{x}, \bar{y})$. Estela apontou-o como um elemento da reta de regressão; Lia reforçou que tal ponto poderia servir para confirmar a tendência global dos dados, quando representado num diagrama de dispersão.

A análise gráfica da correlação foi realizada tendo em vista uma descrição geral da tendência global dos dados (de crescimento ou decrescimento), sem ser ponderada a existência de *outliers* ou a resistência do coeficiente de correlação devido à existência de grupos de dados ou *outliers*, não indo ao encontro das recomendações de Garfield e Ben-Zvi (2008). Por exemplo, uma correlação positiva entre duas variáveis quantitativas foi sempre traduzida pelas duas professoras como *assim que uma aumenta/diminui (respetivamente) a outra também aumenta/diminui (respetivamente)*, mesmo quando se tratavam de distribuições que incluíam alguns dados que não seguiam essa tendência. Este aspeto foi apontado por Engel e Sedlmeier (2011) como sendo particularmente crítico. Cobb et al. (2003) e Garfield e Ben-Zvi (2008) também insistem que as análises de variabilidade nos dados que considerem aspetos locais e globais, bem como o recurso a descrições mais precisas sobre os dados examinados, favorecem o desenvolvimento do raciocínio estatístico dos alunos.

Quanto ao modelo de regressão linear, não foram feitas considerações à sua relevância e ao modo como deve ser usado para estimar valores (Estepa & Batanero, 1996; Sedlmeier & Biehler, 2011). Por exemplo, a leitura da correlação através de diagramas de dispersão pode levar os alunos à ideia de que encontrar o valor do coeficiente de correlação por si só é suficiente tanto para tirar conclusões sobre a existência de relação linear (uma vez que se descurou a forma das distribuições em termos da existência de grupos ou *outliers* e ainda como estes valores podem alterar o valor do coeficiente de correlação) como para aplicar a reta de regressão na previsão de eventos. Estas ocorrências parecem também retratar a leitura que as professoras fizeram ao programa escolar relativamente ao tópico.

Neste tópico, na aula, não houve a intenção de se explorar as ideias dos alunos em torno da reta de regressão e os métodos que estes lhe associaram. O estudo de Sorto et al. (2013) alerta para algumas conceções distintas que os alunos possam ter da *reta que melhor se ajusta aos pontos do diagrama de dispersão*, que podem não ir ao encontro do método que está subentendido nessa expressão. No entanto, essas conceções podem constituir um bom ponto de partida para ajudá-los a chegar aos raciocínios adequados. Assim, os autores sugerem que é importante questionar os alunos sobre as suas estratégias e entender os seus raciocínios para que seja possível orientá-los devidamente para o método que fornece o melhor ajustamento, ou seja, o dos mínimos quadrados. Estela não apresentou uma explicação que convencesse os alunos de que a reta de regressão que se obtém na calculadora gráfica é mais precisa do que qualquer outra reta que se indique como próxima dos dados. Esta situação prendeu-se com o facto de não desejar esquadrinhar a fundo a noção de reta de regressão, à semelhança de Lia, devido, em parte, à sua leitura do programa escolar nesta temática. Contudo, Estela reconheceu a utilidade de uma análise residual no esclarecimento da situação em causa. Adicionalmente, na aula de Estela surgiram combinações de representações no ecrã da calculadora gráfica, algumas das quais resultado de explorações autónomas realizadas pelos alunos. No entanto, não foi privilegiado o aprofundamento dessas situações com leituras apropriadas aos dados e com a promoção de conexões entre essas representações utilizadas, com vista ao estudo da variabilidade e de intuições corretas sobre os dados bivariados.

Nenhuma das professoras valorizou o desenvolvimento do raciocínio estatístico dos alunos no que à previsão de eventos diz respeito. Em particular, aspetos como o *raciocínio com os modelos*, incluindo propriedades e limitações, e a *integração da Estatística e contexto* não foram contemplados pelas professoras. Lia revelou desconhecer algumas propriedades da reta de regressão, em particular, como é que o modelo linear descreve a variabilidade dos dados bivariados, dificuldade também apontada por Engel e Sedlmeier (2011). De uma maneira geral, na exploração das tarefas, nas aulas das duas professoras, não foi evidente a interpretação dos resultados à luz dos dados e do seu contexto (Wild & Pfannkuch, 1999). Ao invés, privilegiou-se a obtenção dos resultados através da calculadora gráfica, o que também limitou o acesso ao raciocínio dos alunos e possivelmente a identificação de dificuldades. Mesmo assim, Estela mostrou mais destreza no uso da calculadora gráfica na exploração deste tópico do que Lia, e acabou por proporcionar uma maior participação dos seus alunos, também impulsionada pelo facto destes demonstrarem já alguma familiaridade com essa ferramenta.

CA. Para as professoras, este tópico é, de uma maneira geral, acessível para os alunos uma vez que os assuntos principais (correlação linear e reta de regressão) podem ser obtidos na calculadora gráfica e a relação bivariada estudada é basicamente a que é traduzida linearmente, sendo, por isso, bastante familiar para os alunos. Caso tivesse mais tempo para explorar o tópico, Estela referiu que teria permitido que os alunos explorassem os dados recolhidos na aula mais calmamente e da forma que quisessem. Lia mencionou que procuraria realizar uma atividade de cariz investigativo com dados bivariados que percorresse a fase de formulação de uma questão sobre uma situação, planificação e execução da recolha de dados, análise de dados e comunicação de resultados.

As professoras foram ajudando os alunos a superar os obstáculos que iam encontrando na utilização da calculadora gráfica como apoio à aprendizagem deste tópico. Estela assinalou um melhor domínio, por parte dos alunos, dos procedimentos da calculadora facilitando a obtenção (rápida) de resultados e representações estatísticas. Por seu turno, Lia notou que os alunos não estavam muito à vontade com a calculadora gráfica, reconhecendo que não lhes tinha dado muitas oportunidades para a experimentar na aula.

Ainda assim, Estela também acabou por deparar-se com várias questões relacionadas com a utilização da calculadora gráfica para se obter as representações e os cálculos solicitados nas tarefas propostas. Neste contexto, houve um maior acesso a dificuldades ou comentários feitos pelos alunos na prática de Estela do que na de Lia, mas não tanto aos seus raciocínios.

CE. O estudo permite destacar alguns aspetos deste domínio quanto à compreensão dos conceitos, representações e dados, e quanto às relações envolvendo conceitos e representações (consideração da *variabilidade*, utilização do modelo linear e realização de previsões). As duas professoras revelam conhecer que a reta de regressão diz respeito à que é obtida através do método dos mínimos quadrados (embora a explicação do fundamento do método não surgisse na aula), e que o centro de gravidade é um ponto que lhe pertence (contudo, Lia não abordou esse aspeto na aula). As professoras descreveram a relação bivariada, em várias situações, sobretudo associada à dependência funcional, não atendendo à variação dos dados (Engel & Sedlmeier, 2011) e não contemplando a resistência do coeficiente de correlação face a *outliers* ou a grupos de dados, nomeadamente, na análise dos dados representados graficamente. Associaram com facilidade, embora de uma forma pouco profunda, a reta de regressão ao coeficiente de correlação e à tendência global de crescimento/decrescimento dos dados. Não deram importância ao centro de gravidade, e também não estabeleceram relação entre este e outros conteúdos e mostraram desconhecer alguns aspetos estruturantes da reta de regressão de forma a aplicarem-na adequadamente na modelação de dados. Relativamente a este assunto, Engel e Sedlmeier (2011) referem que o conhecimento de como a reta de regressão é formada (com as suas componentes estrutural e residual) permite o acesso a propriedades que melhor descrevem os dados e sua variabilidade e ao entendimento do papel da relação bivariada no modelo de regressão linear.

Enquanto Estela revelou ter algumas dificuldades em *raciocinar com o modelo* linear interpretando-o à luz dos *dados e do seu contexto* (Wild & Pfannkuch, 1999), Lia não procurou fazer esse raciocínio. Na realização das previsões com a reta de regressão, Estela sugeriu algumas estratégias para as efetuar na calculadora gráfica. Por seu turno Lia demonstrou não estar tão à vontade, solicitando que os resultados procurados fossem calculados manualmente. Os dados evidenciam algumas dificuldades que as professoras enfrentaram no ensino deste tópico e que condicionaram o CETE e que podem ter limitado, de algum modo, as aprendizagens dos alunos.

7.2.2. Aspetos transversais do conhecimento didático em Estatística

A segunda questão de investigação, *Que aspetos transversais do conhecimento didático em Estatística das professoras se destacam na articulação entre os tópicos do programa?*, é respondida de seguida. À semelhança da secção anterior, esta também é organizada à luz das dimensões do conhecimento didático do professor em Estatística, assumidas no estudo, pela seguinte ordem: conhecimento do ensino do tópico/tema estatístico (*CETE*), conhecimento do aluno (*CA*), conhecimento de Estatística (*CE)* e conhecimento do currículo (*CC).*

O estudo realça alguns aspetos transversais do conhecimento didático do professor em Estatística, na articulação entre os tópicos do programa, que estão relacionados com o desenvolvimento de conexões entre conceitos, representações e dados, no contexto dos assuntos trabalhados nas aulas das duas professoras. Em linhas gerais, quanto à articulação entre os tópicos do programa, o tópico *população e amostra* permite introduzir a matéria-prima trabalhada na Estatística (os conjuntos de dados e seus respetivos contextos) e fazer referência a alguns métodos de recolha de dados, estando, portanto, explicitamente relacionado com todos os outros tópicos cobertos no tema da Estatística, do programa de Matemática A do ensino secundário (ME, 2001). *Organização e interpretação de dados*, *medidas de localização* e *medidas de dispersão* são tópicos relevantes para serem usados num contexto de análise de dados univariados e de comunicação de resultados, uma vez que fazem alusão a algumas representações tabulares e gráficas de dados e medidas estatísticas de centro e de dispersão como síntese de dados. O tópico *dados bivariados e regressão linear* revela-se como o mais complexo do programa escolar por envolver análises de dados bidimensionais, introdução de conceitos e representações numéricas e gráficas que permitem analisar devidamente esses dados bidimensionais. Contudo, este tópico não deixa de estar relacionado com os conceitos e representações numéricas, tabulares e gráficas considerados originalmente para os dados univariados e envolve, ainda, a possibilidade de se realizar extrapolações com os dados bidimensionais. A análise e interpretação dos dados bivariados são mais bem conseguidas quando complementadas com análises individuais a cada dimensão desses dados e quando recorrem a representações variadas com vista a um melhor entendimento dos dados (Garfield & Ben-Zvi, 2008).

CETE. Na aula de Estatística, o professor tem um papel importantíssimo em proporcionar aos alunos a exploração de situações diversificadas que tenham como propósito desenvolver conceitos, suas propriedades e suas relações, e também melhorar gradualmente o desempenho dos alunos na análise de dados e na comunicação dos resultados, com o apoio da tecnologia (Burrill, 2008; Garfield & Ben-Zvi, 2008). Uma experiência deste tipo pode facilitar o desenvolvimento da capacidade de aplicar um conjunto de estratégias que permitam desencadear análises de dados substanciais, complementadas com uma comunicação adequada dos resultados (Franklin et al., 2007; Shaughnessy et al., 2009). Por exemplo, no sentido de desenvolver nos alunos um conhecimento estatístico mais profundo a partir dos gráficos, Curcio e Artzt (1996) e Shaughnessy (2007) indicam que há que libertá-los dos aspetos relacionados com a construção de gráficos e centrá-los na interpretação, análise e extrapolação dos dados referentes a situações reais. Para isso o professor tem de estar consciente da existência dessas situações diversificadas e do potencial de cada uma delas na aprendizagem, tem de saber como é que os conteúdos envolvidos devem ser relacionados, bem como quais as ideias e explicações adequadas para cada situação, indo preferencialmente das mais simples às mais complexas (Garfield & Ben-Zvi, 2008; Groth, 2013; Shaughnessy et al., 2009).

No caso de Estela observou-se que na sua aula, no tema Estatística, surgiram representações diversificadas (numéricas, tabulares e gráficas), maioritariamente obtidas através da calculadora gráfica. Porém, a professora poderia ter tirado mais partido dessas representações, nomeadamente, para analisar propriedades específicas dos conceitos trabalhados, para o estabelecimento de algumas relações, em particular entre as ideias de centro e de dispersão (Burrill, 2008; Garfield & Ben-Zvi, 2008) e também para o entendimento dos dados, ou seja, dando-se atenção a aspetos do raciocínio estatístico tais como consideração da *variabilidade* e *integração de Estatística e contexto* (Wild & Pfannkuch, 1999).

Ao longo do tema, foi também desenvolvida a um nível básico a capacidade de *transnumerar* nas análises de dados, nas quais houve sobretudo um enfoque na obtenção das representações na calculadora gráfica, recorrendo-se pouco à interpretação de dados ou de informações retiradas dessas representações à luz dos dados e do seu contexto. De entre os cinco tópicos trabalhados na aula, Estela deu um pouco mais de destaque à interpretação dos dados no tópico *dados bivariados e regressão linear* nas situações de modelação. Os dados também sugerem que a atuação de Estela no ensino do tema foi bastante influenciada pela sua leitura do programa escolar.

Quanto a Lia, verifica-se que teve o cuidado de promover algumas conexões entre centro e dispersão, envolvendo, por vezes, algumas representações gráficas, sempre com o duplo objetivo de aprofundamento dessas noções e de entendimento dos dados, principalmente nos tópicos *medidas de localização* e *medidas de dispersão*. Assim, é também evidente que, nestes tópicos, Lia integrou de um modo consistente aspetos do raciocínio estatístico na aula, tais como *dados e contexto*, alguma consideração da *variabilidade* e *integração de Estatística e contexto* (Burgess, 2006; Wild & Pfannkuch, 1999). Ainda no caso desta professora, há indícios de uma maior dificuldade do que Estela na conexão de assuntos (entre si e com os assuntos prévios mencionados no tema) no último tópico *dados bivariados e regressão linear*, que parece ser influenciada, em parte, pelo pouco tempo disponível para dedicar a esse tópico. Assim, nesse período, teve de despender tempo a explicar como usar a calculadora gráfica (dado que até iniciar este tópico a única experiência na aula com o menu estatística consistiu na obtenção do valor do desvio-padrão) que era a única forma de obter os resultados de medidas estatísticas e os valores dos coeficientes da reta de regressão, com as ferramentas tecnológicas disponíveis na aula.

Para Lia, a calculadora gráfica podia servir para realizar os cálculos básicos durante a exploração das tarefas mas devia sobretudo ser usada depois de os alunos saberem fazer por si só, manualmente, os cálculos associados às medidas estatísticas e também saberem fazer as representações à mão. Adicionalmente, a professora parece não reconhecer o potencial do uso da tecnologia no ensino do tema para suportar um ensino eficiente ao nível do ensino secundário, tal como autores como Burrill (2008), Franklin et al. (2007) e Garfield e Ben-Zvi (2008) reivindicam.

CA. Para as professoras participantes no estudo, as noções e representações abordadas no tema da Estatística do 10.º ano devem fazer parte do conhecimento do aluno. As capacidades mais trabalhadas pelos alunos nesse tema são efetivamente a realização de cálculos e a construção ou determinação de representações gráficas, tal como o estudo de Quintas et al. (2009) sugere. Ambas consideram que a interpretação de um resultado e a realização de uma investigação estatística são atividades difíceis para os alunos e completamente fora do habitual numa aula de Matemática. Daqui pode-se também inferir que a compreensão dos dados não parece ser reconhecida pelas professoras como um conhecimento crucial para os alunos, ao contrário das recomendações de vários autores (por exemplo, Shaughnessy, 2007).

Lia insistiu que os alunos realizassem alguma interpretação à luz dos dados e contexto através de algumas análises comparativas de dados sobretudo nos tópicos *Medidas de localização* e *medidas de dispersão*. Denotando dificuldades dos alunos em produzir comentários nessas situações, teve o cuidado de completar constantemente as suas respostas. Ao longo do tema, nos alunos de Lia verificaram-se dificuldades no entendimento de situações e resolução das tarefas que incorporavam dados agregados em classes. Os alunos, de uma maneira geral, não conseguiam fazer uma leitura de tabelas ou gráficos de modo a ficarem a saber o significado dos dados, bloqueando logo à partida as suas respostas às questões colocadas. Contudo, esta situação não parece ser alheia ao facto de, quando os alunos tiveram os primeiros contactos com dados agregados (no tópico *Organização e interpretação dos dados*), ter sido sobretudo privilegiada a construção tabular ou gráfica. Complementar essas experiências com leituras de dados significativas em representações tabulares ou gráficas é uma forma eficiente de ajudar os alunos a apropriarem-se adequadamente destes conhecimentos (Curcio & Artzt, 1996; Shaughnessy, 2007).

Na aula de Estela ocorreram várias situações em que poderia ter promovido mais conexões entre as representações obtidas na calculadora gráfica com o objetivo de alcançar um aprofundamento dos conceitos estudados e também o entendimento dos dados ao longo do tópico. Aliás, só no último tópico é que os alunos desta professora começaram a demonstrar alguma destreza na utilização do *menu* da Estatística das calculadoras gráficas que utilizaram e, simultaneamente, alguma capacidade para comentar uma ou outra representação obtida, mas a pedido da professora. Assim, ao longo do tema, Estela procurou ajudar os alunos especialmente a superar as dificuldades com que se deparavam na utilização da calculadora gráfica durante a resolução das tarefas propostas, sendo as mais frequentes as que tinham a ver com obter o resultado certo ou aceder a uma determinada representação, muitas das vezes, por falharem nalgum passo da instrução que deviam seguir. Como os alunos de Lia tiveram poucas oportunidades para usar o *menu* Estatística da calculadora gráfica no tema, no tópico final acabaram por revelar ainda mais dificuldades do que os de Estela em obter os resultados e as representações apropriadas com essa tecnologia.

Verifica-se, pois, a importância do conhecimento do professor acerca das dificuldades dos alunos no tema da Estatística, assim como a necessidade de um maior acesso aos seus raciocínios, nomeadamente quando estes trabalham com situações mais desafiantes.

CE. Relativamente aos conhecimentos de Estatística, verifica-se que há necessidade de as professoras desenvolverem ideias variadas dos conceitos abordados nas aulas, conhecer e lidar com propriedades destes conceitos, saber produzir afirmações adequadas relativamente à variabilidade dos dados, em diferentes circunstâncias, que envolvam representações gráficas e numéricas diversas, e em raciocinar sobre modelos. Adicionalmente, no alargamento deste conhecimento a tecnologia pode ter um papel preponderante (Garfield & Ben-Zvi, 2008). Efetivamente, Estela revelou, de uma maneira geral, um maior domínio das potencialidades da calculadora gráfica do que Lia. Contudo, em várias instâncias, as duas professoras poderiam ter tirado mais proveito desta tecnologia para estabelecer conexões entre representações mais profundas.

CC. As professoras consideram que a Estatística é um tema relevante na formação dos alunos e que, por isso, deve integrar o programa de Matemática A, o que vai ao encontro do estudo de Quintas et al. (2009). O conhecimento curricular das professoras no tema é bastante semelhante, na medida que consideram que o programa escolar cobre a Estatística descritiva, mas não a inferencial, e assim solicita que se trabalhe na aula este tema com foco nos cálculos e na construção e determinação das representações, na qual a interpretação dos resultados geralmente não é contemplada. Adicionalmente, nas leituras ao programa escolar que fizeram tópico a tópico no tema, observou-se também a tendência para salientarem aspetos processuais com respeito às noções e representações contempladas, descurando algumas ligações relevantes entre os assuntos estudados. Estes resultados estão em linha com os de outros estudos centrados no professor no âmbito da Estatística (Groth & Bergner, 2006; Watson & Callingham, 2013).

Apesar de as professoras também indicarem que as investigações estatísticas não surgem no programa de Matemática A, em conversa com a investigadora reconheceram-nas como atividades complexas. Uma atividade dessa natureza ao ser proposta na disciplina de Matemática A, no 10.º ano, exigiria na sua opinião a mobilização de todos os conhecimentos estatísticos adquiridos. Desafiadas pelas conversas com a investigadora e pelas reflexões que realizaram ao longo do estudo, Estela indicou que gostaria de aprofundar melhor os vários conceitos estatísticos à custa de uma maior exploração de gráficos diversificados, e Lia que gostaria de enfatizar a interpretação e comunicação dos resultados à luz dos dados e do seu contexto, o que evidencia uma certa ampliação do seu conhecimento curricular.

A dificuldade de mudança de paradigma no ensino da Estatística no sentido de considerar uma aprendizagem mais conceptual e relacional, envolvendo a realização de análises que permitam compreender os dados e dar resposta às questões colocadas sobre esses dados e que inclua a promoção do raciocínio estatístico dos alunos tem vindo a ser salientada por vários autores, em particular, Ben-Zvi e Garfield (2008), Henriques e Ponte (2014), Pfannkuch e Ben-Zvi (2011) e Shaughnessy (2007). Contudo, no contexto escolar português essa dificuldade é muitas vezes agravada por alguns constrangimentos que tendem a influenciar a gestão curricular na Estatística, nomeadamente, o tema Estatística é o que tipicamente fica para o fim do programa e é o que acaba por sofrer mais cortes no tempo que lhe dedicam (Turkman & Ponte, 2000). Em concordância com esta situação, as professoras participantes do estudo expressaram que ao tema Estatística não é atribuída a mesma importância que outros temas do programa de Matemática (ME, 2001), nomeadamente, ao nível da avaliação dos alunos nos exames nacionais do ensino secundário, considerando que esta perspetiva é partilhada pela generalidade dos professores de Matemática. Isto reflete-se nas suas experiências profissionais, levando a que o tema seja habitualmente lecionado de um modo superficial e rápido, o que vai ao encontro dos resultados do estudo de Boaventura e Fernandes (2004).

7.3 Reflexão final

Esta reflexão final incide sobre algumas ideias em torno do conhecimento didático do professor em Estatística, apresenta algumas implicações do estudo para a formação (inicial e contínua) e para o desenvolvimento profissional, e aponta algumas limitações do estudo e, ainda, sugestões para futuras investigações na área da Educação Estatística.

O estudo revela elementos acerca da natureza do conhecimento didático de Estatística que as professoras que participaram no estudo possuem, evidenciando o seu carater dinâmico, situacional, experiencial e pessoal (Canavarro, 2003; Guimarães, 2008; Ponte, 2012) e também multidimensional (Chapman, 2013; Depaepe et al., 2013). Os resultados do estudo também permitem entender algumas relações entre os domínios do conhecimento didático em Estatística das professoras.

Em termos gerais, com esta investigação é possível perceber que o modelo do conhecimento didático de Estatística mobilizado permite a identificação de um conjunto concreto de conhecimentos que enfatizam a especificidade do ensino dos vários assuntos da unidade de Estatística, na disciplina de Matemática A no 10.º ano. A análise realizada com as lentes do conhecimento didático em Estatística permitiu compreender, de uma forma aprofundada, o trabalho desenvolvido pelas duas professoras no ensino dos diferentes tópicos de Estatística, integrando as decisões tomadas pelas professoras e identificando fatores que influenciaram positiva ou negativamente a sua ação.

Os conhecimentos didáticos de Estatística das duas professoras foram também confrontados, tópico a tópico, com vista a salientar diferenças e semelhanças para melhor entendimento do fenómeno em estudo. E, apesar de alguns tópicos estatísticos terem sido desenvolvidos nas aulas das professoras de uma forma tendencialmente compartimentada (por exemplo, *População e amostra* e *Organização e interpretação de dados*), também se procurou analisar as experiências potencialmente significativas que emergiram nestes contextos e caracterizar globalmente os seus conhecimentos didáticos ao longo do tema com elementos que se destacaram na articulação entre os tópicos. Assim, através da análise realizada da prática das professoras foi também possível refletir sobre como desenvolver um ensino da Estatística que vá ao encontro da promoção do raciocínio estatístico dos alunos (Franklin et al., 2007).

Acrescenta-se, ainda, que a natureza interdependente dos domínios do conhecimento didático do professor, por vezes, dificultou a classificação das situações relativas à prática das professoras no domínio adequado mas, por outro lado, também contribuiu para uma maior reflexão sobre o assunto, procurando-se clarificar cuidadosamente quais os aspetos do conhecimento em cada domínio que deveriam ser enfatizados no estudo.

Ao caracterizar a especificidade do conhecimento didático do professor no ensino da Estatística, este estudo pode também dar um contributo importante para a formação inicial e contínua dos professores dos ensinos básico e secundário. O facto de os resultados do estudo se organizarem por tópicos estatísticos, e nas quatro vertentes do conhecimento didático, permite identificar uma multiplicidade de aspetos ou situações que podem ser contempladas num contexto formativo, relativas, por exemplo, a dificuldades na aprendizagem, a lacunas no conhecimento estatístico do professor e à compreensão dos conceitos e dos dados, na condução do ensino. Também o modelo de conhecimento didático que foi adotado no estudo pode informar acerca de elementos importantes a integrar na formação de professores ou futuros professores, no âmbito da didática da Estatística (Groth, 2013). Pelo facto de se tratar de um estudo realizado a partir da prática das professoras, assente no contexto de sala de aula e evidenciando as suas complexidades, pode assim exercer uma maior influência na promoção do conhecimento didático (Depaepe et al., 2013) no domínio da Estatística. Nomeadamente, os resultados do estudo que evidenciam aspetos a melhorar nas diferentes dimensões estudadas e que envolvem exemplos de situações exploradas pelo professor ou pelos alunos, poderão ter grande potencial formativo (Burgess, 2008, 2009).

Tendo em conta os resultados do estudo, verifica-se existir um conjunto de desafios que se levantam ao ensino da Estatística no ensino secundário, principalmente se se desejar alcançar o raciocínio estatístico dos alunos (Batanero et al., 2013; Burrill & Biehler, 2011). O essencial do trabalho do professor no ensino da Estatística não reside apenas em propor nas aulas tarefas diversificadas em cada tópico que tenham potencial para promover o raciocínio estatístico. É também necessário pensar na forma de as explorar de um modo significativo e integrar numa sequência de ensino coerente que envolva a realização de conexões relevantes, e cada vez mais profundas, por meio dos conceitos, representações e dados, dentro de cada tópico e entre tópicos (Ben-Zvi & Garfield, 2004; Garfield & Ben-Zvi, 2008). Sendo igualmente importante que sejam promovidas interações com os alunos com vista à compreensão conceptual, à realização de análises de dados cada vez mais ricas, que incorporem exploração dos dados reais através de representações variadas, eventualmente, suportadas pela tecnologia, e que contemplem o entendimento dos dados, e da sua variabilidade, em cada situação problemática (Ben-

Zvi & Garfield, 2004; Shaughnessy et al., 2009). Um ensino mais profícuo no tema Estatística do ensino secundário passa também pela capacidade do professor aceder e avaliar adequadamente os raciocínios dos alunos, bem como as estratégias que utilizam em diferentes contextos, de modo a regular a sua ação e tomar decisões sobre o ensino. Para tal, o professor deve possuir uma preparação que inclua a compreensão dos conceitos estatísticos chave (Burrill & Biehler, 2011), experiências em lidar com análises de dados e o processo investigativo, de uma forma geral (Batanero et al., 2013; Ben-Zvi & Garfield, 2004; Snee, 1993; Wild & Pfannkuch, 1999) e o desenvolvimento do seu conhecimento didático na área. Outro desafio do ponto de vista do desenvolvimento profissional no domínio da didática da Estatística é o facto das orientações curriculares portuguesas do ensino secundário (MEC, 2014) não enfatizarem o desenvolvimento do raciocínio estatístico dos alunos, quando é amplamente reconhecido o papel da Estatística na formação dos alunos e que cabe à escola providenciar uma aprendizagem significativa nesta área (Franklin et al., 2007; Henriques & Ponte, 2014; Ponte et al., 2006).

Quanto a limitações do estudo, há a considerar o facto de existirem dimensões do conhecimento didático que tiveram uma menor expressão na análise de dados do que outras, tais como o CC e o CE. Apesar de a investigadora ter recorrido a vários métodos de recolha dados para obter o máximo da informação possível para todas essas dimensões, na dimensão CC, as professoras tinham a tendência de caracterizar os tópicos estatísticos e o tema Estatística em linhas gerais. Porém, na aula, foram concretizando de forma diferenciada, algumas conexões dentro dos tópicos e entre tópicos. Para a dimensão CE, a investigadora gostaria de ter reunido mais dados, mas optou por ser mais comedida pois não queria criar constrangimentos nas professoras, no sentido de pensarem que estavam a ser avaliadas quanto aos seus conhecimentos estatísticos. Assim, no estudo, apesar de a dimensão CE incluir uma perspetiva mais geral dos aspetos realçados, permitiu retirar algumas conclusões particulares e também nomear alguns conhecimentos estatísticos relevantes que as professoras possuíam, mas que não mobilizaram nas suas aulas observadas. Outra possível limitação do estudo tem a ver com o número de casos constituídos. A opção pelo estudo de um maior número de casos poderia permitir atingir um conhecimento mais abrangente do conhecimento didático do professor em Estatística. No entanto, o facto de se optar por uma análise em profundidade de cada um deles, num conjunto alargado de tópicos e ideias estatísticas, inviabilizou a realização de um número maior de casos.

Há vários resultados do estudo que poderão ser constituídos como um bom ponto de partida para ser desenvolvidos em investigações futuras sobre o ensino e aprendizagem da Estatística no ensino secundário. Ainda, existem poucos estudos sobre o conhecimento profissional do professor em Estatística no ensino secundário, que prestem uma atenção detida tanto à ação do professor no seu contexto profissional como às suas explicações sobre a sua atuação, e que cubram por completo uma unidade de Estatística, tal como sucedeu neste estudo. De facto, estes resultados podem apoiar investigações que adotem uma perspetiva mais situada do conhecimento didático do professor que ensina a Estatística ou que assumam uma perspetiva mais cognitiva deste conhecimento (isto é, que procuram medir este conhecimento, recorrendo frequentemente a testes) ou, ainda, outras que conjuguem elementos dessas duas perspetivas. Ainda, no que refere a sugestões para futuros estudos, emergem algumas sugestões que se prendem com os alunos e a aprendizagem, nomeadamente, parece relevante que se compreenda de uma forma aprofundada as dificuldades dos alunos quanto à apropriação da calculadora gráfica e de outros recursos tecnológicos (*software*, tais como *Excel*, *TinkerPlots*, através do computador) no tema da Estatística. Também seria interessante compreender-se o potencial do uso da tecnologia no ensino e aprendizagem do tema da Estatística, sobretudo em contextos que visem objetivos de aprendizagem mais ambiciosos, tal como é o caso do raciocínio estatístico. Estes contextos requerem uma participação mais ativa do aluno, exploração de situações diversificadas, uma compreensão conceptual mais aprofundada e também estudo de situações progressivamente mais complexas que integrem análises de variabilidade dos dados reais e comunicação de resultados, suportado pelas representações numéricas e gráficas obtidas com os recursos tecnológicos utilizados e entendimento dos dados e do seu contexto.

REFERÊNCIAS

Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician*, *27*(1), 17-21.

Adler, P. A., & Adler, P. (1994). Observational techniques. In N. K. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 377–392). Thousand Oaks: Sage Publications, Inc.

Almeida, R. (2000). *Imagens sobre o ensino e a aprendizagem da Estatística.* Dissertação de mestrado, Universidade de Lisboa, Lisboa.

APM (1998). *Matemática 2001: Diagnóstico e recomendações para o ensino da matemática: Relatório preliminar*. Lisboa: APM.

Ball, D. (2002). Knowing mathematics for teaching: Relations between research and practice. *Mathematics and Education Reform Newsletter*, *14*(3), 1–5.

Ball, D., Hill, H., & Bass, H. (2005). Knowing mathematics for teaching. Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, *29*(1), 14-17, 20-22, 43-46.

Ball, D., Lubienski, S., & Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 433–456). New York: Macmillan.

Ball D., Thames M., & Phelps G. (2005). Articulating domains of mathematical knowledge of teaching. *Paper presented at the 2005 annual meeting of the American Education Research Association*, Montreal, Canada.

Ball D., Thames M., & Phelps G. (2008). Content knowledge for teaching. What makes it special? *Journal of Teacher Education*, *59*(5), 389–407.

Batanero, C. (2001). *Didáctica de la estadística*. Grupo de Investigación en Educación Estadística, Departamento de Didáctica de la Matemática. Granada: Universidad de Granada.

Batanero, C., & Diaz, C. (2010) Training teachers to teach statistics: What can we learn from research? *Statistique et Enseignement*, *1*(1), 5–20.

Batanero, C., Diaz, C., Contreras, J. M., & Roa, R. (*2013*). El sentido estadístico y su desarrollo. *Números*, *83,* 7–18.

Batanero, C., Garfield, J., Ottaviani, M. G., & Truran, J. (2001). Research in statistics education: Some priority questions. *Statistical Education Research Newsletter, 1*(2), 2-5.

Batanero, C., & Godino, J. (2005). Perspectivas de la educación estadística como área de investigación. In R. Luengo (Ed.), *Líneas de investigacion en Didáctica de las Matemáticas* (pp. 203–226). Badajoz: Universidad de Extremadura.

Batanero, C., Godino, J., & Roa, R. (2004). Training teachers to teach probability. *Journal of Statistics Education, 12*(1).

www.amstat.org/publications/jse/v12n1/batanero.html

Ben-Zvi D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions and challenges. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3–15). New York: Kluwer Academic Publishers.

Beswick, K., Callingham, R., & Watson, J. (2012). The nature and development of middle school mathematics teachers' knowledge. *Journal of Mathematics Teacher Education*, *15*, 131–157.

Biggs J. B. & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. NY: Academic.

Boaventura, M., & Fernandes, J. A. (2004). Dificuldades de alunos do 12.º ano nas medidas de tendência central. O contributo dos manuais escolares. In J. A. Fernandes, M. Sousa & S. Ribeiro (Eds.), *Ensino e aprendizagem de probabilidades e estatística. Atas do I Encontro de Probabilidades e Estatística na Escola* (pp. 103–126). Braga: CIEd da Universidade do Minho.

Branco, J. (2000). Estatística no secundário: O ensino e seus problemas. In C. Loureiro, F. Oliveira, & L. Brunheira (Eds.), *Ensino e aprendizagem da estatística* (pp. 11–30). Lisboa: SPE, APM, DE e EIO da FCUL.

Burgess, T. A. (2006). A framework for examining teacher knowledge as used in action while teaching statisticcs. In A. Rossman & B. Chance (Eds.), *Work cooperatively* *in Statistics Education: Proceedings of the seventh International Conference on Teaching Statistics*. Salvador, Brazil.

Burgess, T. A. (2008). Teacher knowledge for teaching statistics through investigations. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* Monterrey, Mexico.

Burgess, T. (2009). Teacher knowledge and statistics: What types of knowledge are used in the primary classroom? *The Montana Mathematics Enthusiast*, *6*(1-2), 3–24.

Burgess, T. (2011). Teacher knowledge of and for statistical investigations. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education. A joint ICMI/IASE study: The 18th ICMI Study* (pp. 259–270). NY: Springer.

Burrill, G. (2008). Fundamental ideas in teaching statistics and how they affect the training of teachers. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* Monterrey, Mexico.

Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education. A joint ICMI/IASE study: The 18th ICMI Study* (pp. 57–70). NY: Springer.

Bogdan, R. & Biklen, S. K. (1994). *Investigação qualitativa em educação. Uma introdução à teoria e aos métodos*. Porto: Porto Editora.

Callingham, R. (1997). Teachers' multimodal functioning in relation to the concept of average. *Mathematics Education Research Journal*, *9*(2), 205–224.

Callingham, R. & Watson, J. (2011). Measuring levels of statistical pedagogically content knowledge. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education. A joint ICMI/IASE study: The 18th ICMI Study* (pp. 283-293). NY: Springer.

Canavarro, A. P. (2003). *Práticas de ensino de Matemática: Duas professoras, dois currículos*. Dissertação de doutoramento, Lisboa: APM.

Carvalho, C. (2001). *Interações entre pares: Contributos para a promoção do desenvolvimento lógico e do desempenho estatístico, no 7º ano de escolaridade*. Dissertação de doutoramento, Lisboa: APM.

Chapman, O. (2013). Investigating teachers' knowledge for teaching mathematics. *Journal of Mathematics Teacher Education*, *16*, 237–243.

Cobb, G., & Moore, D. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, *104*(9), 801–823.

Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction*, *21*(1), 1–78.

Cockcroft, W. H. (1982). *Mathematics counts*. London: HMSO.

Coolidge, F. L. (2013). *Statistics: a gentle introduction*. Sage Publications.

Costa, B., & Rodrigues, E. (2010). *Novo Espaço*. *Matemática A, 10.º ano*. Porto Editora.

Curcio, F., & Artzt, A. (1996). Assessing students' ability to analyze data: Reaching beyond computation. *Mathematics Teacher 89*, 668–673.

DelMas, R. (2002). Statistical literacy, reasoning and thinking: A commentary. *Journal of Statistical Education*, *10*(3). www.amstat.org/publications/jse/

DelMas, R. (2004). A comparison on mathematical reasoning and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 79–95). New York: Kluwer Academic Publishers.

DelMas, R. & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal*, *4*(1), 55–82.

Denzin, N., & Lincoln, Y. (1998). *Strategies of qualitative inquiry*. Sage publications.

Depaepe, F., Torbeyns, J., Vermeersch, N. Janssens, D., Janssens, R., Kelchtermans, G., Verschaffel, L., & Dooren, W. (2015). Teachers' content and pedagogical content knowledge on rational numbers: A comparison of prospective elementary and lower secondary school teachers. *Teaching and Teacher Education*, *47*, 82–92.

Depaepe, F., Verschaffel, L, & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, *34*, 12-25.

Duarte, J. A. (2011). *Tecnologias e pensamento algébrico: Um estudo sobre o conhecimento profissional dos professores de Matemática*. Dissertação de doutoramento, Instituto de Educação, Universidade de Lisboa, Lisboa.

Engel, J., & Sedlmeier (2011). Correlation and regression in the training of teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education. A joint ICMI/IASE study: The 18th ICMI Study* (pp. 247–258). New York, NY: Springer.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119–161). London: MacMillan Publishing Company.

Estepa, A., & Batanero, C. (1996). Judgments of correlation in scatter plots: students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education*, *4*, 25–41.

Estepa, A., & Cobo, F. (1998). Correlation and regression in secondary school textbooks. *ICOTS 5*, 672–677. http://iase-web.org/

Fernandes, J. (2009). Ensino e aprendizagem da Estatística: Realidades e desafios. *Atas do XIX Encontro de Investigação em Educação Matemática*. *Números e Estatística - Refletindo no presente, perspetivando o futuro*. Vila Real, Portugal: SPCE. http://www.spiem.pt/publicacoes/arquivo/encontro-2009/

Fernandes, J., Carvalho, C, & Correia, P. (2011). Contributos para a caracterização do ensino da Estatística nas escolas. *Bolema*, *24*(39), 585–606.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report. A Pre-k-12 Curriculum Framework*. Alexandria, VA: ASA.

Friel, S. (1998). Teaching statistics: What's average? In L. Morrow & M. Kenney (Eds.), *The teaching and learning of algorithms in school mathematics* (pp. 208–217). Reston, VA: NCTM.

Friel, S., Curcio, F., & Bright, G. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, *32*(2), 124–158.

Friel, S., Connor, W., & Mamer, J. (2006). More than "Meanmedianmode" and a bar graph: What's needed to have a statistical conversation. In G. Burrill (Ed.) *Thinking and reasoning with data and chance: Sixty-eighth yearbook of the NCTM* (pp. 117-138). Reston, VA: NCTM.

Garfield, J., & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, *4*(1), 92–99. http://www.stat.auckland.ac.nz/serj.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, *75*(3), 372–396.

Garfield, J. & Ben-Zvi, D. (2008). *Developing students' statistical reasoning-connecting research and teaching practice*. Springer Science+Business Media B. V.

Gattuso, L., & Ottaviani, M. G. (2011). Complementing mathematical thinking and statistical thinking in school mathematics. In C. Batanero, G. Burrill & C. Reading (Eds.) *Teaching statistics in school mathematics – Challenges for teaching and teacher education. A joint ICMI/IASE study: The 18th ICMI Study* (pp. 121–132). New York, NY: Springer.

Godino, J., Batanero, C., Roa, R., & Wilhelmi, M. (2008). Assessing and developing pedagogical content and statistical knowledge of primary school teachers through project work. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Joint ICMI/IASE study: Teaching statistics in school mathematics. Challenges for teaching and teacher education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* Monterrey, Mexico.

González, O. (2013). Conceptualizing and assessing secondary mathematics teachers' professional competencies for effective teaching of variability-related ideas. In B. Burz, C. Haser & M. Mariotti (Eds.), *Proceedings of the 8th Congress of the European Research in Mathematics Education* (pp. 809–818). Ankara, Turkey: Middle East Technical University.

Groth, R. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education, 38*(5), 427–437.

Groth, R. (2009). Characteristics of teachers' conversations about teaching mean, median and mode. *Teaching and Teacher Education*, *25*(5), 707-716.

Groth, R. (2013). Characterizing key development understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework. *Mathematical Thinking and Learning*, *15*(2), 121–145.

Groth, R., & Bergner, J. (2005). Pre-service elementary school teachers' metaphors for the concept of statistical sample. *Statistics Education Research Journal*, *4*(2), 27–42.

Groth, R., & Bergner, J. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning*, *8*, 37–63.

Guimarães, F. (1999). O conteúdo do conhecimento profissional de duas professoras de Matemática. *Quadrante*, *8*, 5–32.

Guimarães, H. M. (2008). Perspetivas sobre o conhecimento do professor. *Revista Diálogo Educacional*, *8*(25), 819–839.

Guimarães, H. M. (2010). Sebastião e Silva e as perspetivas metodológicas para o ensino da Matemática em Royaumont (1959). In H. Gomes, L. Menezes & I. Cabrita (Eds.), *Atas do XXI SIEM* (741-752). Aveiro, Portugal. Lisboa: APM.

Hardiman, P., Pollatsek, A., & Well, A. (1986). Learning to understand the balance beam. *Cognition and Instruction*, *3*(1), 63–86.

Henriques, A., & Oliveira, H. (2013). Prospective teacher's statistical knowledge for teaching when analysing classroom episodes. In A. M. Lindmeier & A. Heinze

(Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (v. 3, pp. 41–48). Kiel, Germany: PME.

Henriques, A., & Ponte, J. P. (2014). Preparing teachers to teach statistics: developing professional knowledge and practice. In K. Makar, B. de Sousa & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics* (ICOTS9). Flagstaff, Arizona, USA. Vooburg, The Netherlands: ISI.

Hill, H., Rowan, B., & Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Education Research Journal*, *42*(2), 371–406.

Hill, H., Schilling, S., & Ball, D. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, *105*, 11–30.

Howell. H. (2012). *Charactering mathematical knowledge for secondary teaching: A case from high school algebra*. PhD. Dissertation, New York University, ProQuest LLC.

Huberman, A. M., & Miles, M. B. (1994). Data management and analysis methods. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 428–444). Thousand Oaks, California: Sage publications.

Konold, C., & Pollatsek, A. (2004). Data analysis as the search for signals in noise processes. *Journal for Research in Mathematics Education*, *33*(4), 259–289.

Langrall, C., & Mooney, E. (2010). Probability and Statistics in secondary school classrooms. *Teaching and Learning Mathematics*, 27–31.

Lessard-Hébert, M., Goyette, G., & Boutin, G. (1994). *Investigação qualitativa: Fundamentos e práticas*. Lisboa: Instituto Piaget.

Ludke, M., & André, M. (1986). *Pesquisa em educação: Abordagens qualitativas*. São Paulo. EPU.

Ma, L. (1999). *Knowing and teaching elementary mathematics: teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Makar, K. & Confrey, J. (2005). "Variation-talk": articulating meaning in statistics. *Statistics Education Research Journal*, *4*(1), 27–54. http://stat.auckland.ac.nz/serj

Marshall, C., & Rossman, G. (2006). *Designing qualitative research*. Sage Publications.

Martins, M. E. (2012). Dados são mais do que números. *Educação e Matemática*, *120*, 37–41.

ME (1997). *Matemática. Programas do 10.º, 11.º e 12.º anos*. Lisboa: Editorial do Ministério da Educação.

ME (2001). *Programa de Matemática A, 10º ano*. Lisboa: Editorial do Ministério da Educação.

ME (2007). *Programa de Matemática do ensino básico*. Lisboa: DGIDC.

MEC (2014). *Programa e metas curriculares de Matemática A para o ensino secundário*. Lisboa: MEC. http://www.dge.mec.pt/

Merriam, S. (1988). *Case study research in education: A qualitative approach*. Jossey-Bass Inc., Publishers.

Moore, D. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association*, *93*, 1253–1259.

Moritz, J. (2004). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227–256). New York: Kluwer Academic Publishers.

Munby, H., Russell, T., & Martin, A. (2001). Teachers' knowledge and how it develops. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 877–904). Washington: AERA.

NCTM (2007). *Princípios e normas para a matemática escolar*. Lisboa: APM.

Oliveira, H. (1998). *Atividades de investigação na aula de Matemática: Aspetos da prática do professor*. Dissertação de mestrado, Departamento de Educação da Faculdade de Ciências, Universidade de Lisboa, Lisboa.

Oliveira, H., & Henriques, A. (2014). Um quadro de análise do conhecimento estatístico para ensinar de futuros professores. *Boletim Gepem* (online), n. 64, 2176-2988.

Oliveira, H., & Ponte, J. P. (1997). Investigação sobre conceções, saberes e desenvolvimento profissional de professores de Matemática. *Atas do SIEM VII* (pp. 3–23). Lisboa: APM.

Ottaviani, M. G. (1987). Some notes on a history of teaching Statistics in Italy. *Statistica*, *47*(4), 619-648.

Ottaviani, M. G. (1998). Developments and perspectives in statistical education. *Proceedings of the joint IASS/IAOS Conference Statistics for Economics and Social Development. Mexico.*

Ottaviani, M. G. (2002). Statistics, from a tool for State and Society to a tool for all citizens. *International Statistical Review*, *70*(1), 30–22.

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. California: Sage Publications.

Pestana, D., & Velosa, S. (2010). *Introdução à Probabilidade e à Estatística*, v. 1. Lisboa: Fundação Calouste Gulbenkian.

Petrou, M., & Goulding, M. (2011). Conceptualising teachers' mathematical knowledge in teaching. In T. Rowland & K. Ruthven (Eds.) *Mathematical knowledge in teaching* (v. 50, pp. 9–24). London: Springer

Pollatsek, A., Lima, S., & Well, A. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics*, *12*, 191–204.

Ponte, J. P. (1998). Da formação ao desenvolvimento profissional. *Atas do ProfMat 98* (pp. 27–44). Lisboa: APM.

Ponte, J. P. (1999). Didáticas específicas e construção do conhecimento profissional. In J. Tavares, A. Pereira, A. Pedro & H. Sá (Eds.), *Investigar e formar em educação: Atas do congresso da SPCE* (pp. 59–72). Porto: SPCE.

Ponte, J. P. (2006). Estudos de caso em educação matemática. *Bolema*, *25*, 105–132.

Ponte, J. P. (2008). A investigação em educação matemática em Portugal: Realizações e perspetivas. In R. Luengo-González, B. Gómez-Alfonso, M. Camacho-Machín, & L. Nieto (Eds.), *Investigación en educación matemática XII* (pp. 55–78). Badajoz: SEIEM.

Ponte, J. P. (2011). Preparing teachers to meet the challenges of statistics education. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education. A joint ICMI/IASE study: The 18th ICMI Study* (pp. 299–310). New York, NY: Springer.

Ponte, J. P. (2012). Estudando o conhecimento e o desenvolvimento profissional do professor de matemática. In N. Planas (Ed.), *Educación matemática: Teoría, critica y prática* (pp. 83–98). Barcelona: Graó.

Ponte, J. P. (2014). Formação do professor de Matemática: Perspetivas atuais. In J. P. Ponte (Ed.), *Práticas profissionais dos professores de matemática* (pp. 344–358). Lisboa: IEUL.

Ponte, J. P., Brocardo, J. & Oliveira, H. (2006). *Investigações matemáticas na sala de aula*. Autêntica Editora. Belo Horizonte.

Ponte, J. P., Boavida, A., Graça, M., & Abrantes, P. (1997). *Didática da Matemática*. Lisboa: ME, Departamento do Ensino Básico.

Ponte, J. P., & Chapman, O. (2006). Mathematics teachers' knowledge and practices. In A. Gutiérrez & P. Boeno (Eds.), *Handbook of research on the psychology of mathematics education: Past, present, and future* (pp. 461–494). Roterdham: Sense.

Ponte, J. P., & Fonseca, H. (2001). Orientações curriculares para o ensino da estatística: Análise comparativa de três países. *Quadrante*, *10*(1), 93–115.

Ponte, J. P., & Oliveira, H. (2002). Remar contra a maré: A construção do conhecimento e da identidade profissional na formação inicial. *Revista de Educação*, *11*(2), 145–163.

Ponte, J. P., & Santos, L. (1998). Práticas letivas num contexto de reforma curricular. *Quadrante*, *7*(1), 3–33.

Ponte, J. P., & Turkman, M. A. (2000). Introdução. In C. Loureiro, F. Oliveira & L. Brunheira (Eds.), *Ensino e aprendizagem da estatística* (pp. 5–9). Lisboa: SPE, APM, DE e EIO da FCUL.

Quintas, S., Oliveira, H., & Ferreira, R. (2009). Estudo exploratório sobre perspetivas e práticas de professores de Matemática, do ensino secundário, no domínio do ensino da Estatística. *Atas do XIX EIEM*. Vila Real, Portugal.

Quintas, S., Oliveira, H., & Ferreira, R. (2010). Secondary mathematics teachers' perspectives and practices in Statistics. In M. Pinto & T. Kawasaki (Eds.), *Proceedings of the Annual Conference of the International Group for the Psychology of Mathematics Education* (pp. 91–92). Brazil: PME.

Reaburn, R. (2012). Strategies used by students to compare two data sets. In J. Dindyal, L.P. Cheng & S. F. Ng (Eds.), *Mathematics education: Expanding horizons. Proceedings of the 35th annual conference of the mathematics education research group of Australasia* (pp. 633–639). Singapore: MERGA.

Reading, C., & Shaughnessy, M. (2004). Reasoning about variation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201–226). New York: Kluwer Academic Publishers.

Ribeiro, M., Carrillo, J., & Monteiro, R. (2009). O conhecimento profissional em ação durante a exploração de um pictograma: uma situação de (i) literacia. *Atas do XIX EIEM: Números e Estatística- refletindo no presente, perspetivando o futuro*. Vila Real, Portugal: SPCE. http://www.spiem.pt/publicacoes/arquivo/encontro-2009/

Rossman, A., Chance, B., & Medina, E. (2006). Some important comparisons between Statistics and Mathematics, and why teachers should care. In G. Burrill (Ed.) *Thinking and reasoning with data and chance: Sixty-eighth yearbook of the NCTM* (pp. 323-334). Reston, VA: NCTM.

Rowland, T. (2013). The knowledge quartet: A framework for analysing and developing mathematics teaching. In J. A. Fernandes, H. Martinho, J. Tinoco & F. Viseu. (Eds.), *Atas do XXIV Seminário de Investigação em Educação Matemática* (pp. 31–47). Braga: CIE da Universidade do Minho.

Rowland, T., Huckstep, P., & Thwaites, A. (2005). Elementary teachers' mathematics subject knowledge: The knowledge quartet and the case of Naomi. *Journal of Mathematics Teacher Education*, *8*, 255–281.

Rowland, T., Turney, F., & Thwaites, A. (2013). Developing mathematics teacher education practice as a consequence of research. In B. Ubuz, C. Haser & M.

Mariotti (Eds.) *Proceedings of the Eighth Congress of the European society for Research in Mathematics Education* (pp. 3227-3236). Ankara, Turkey: Middle East Technical University.

Rowland, T., Turney, F., Thwaites, A., & Huckstep, P. (2009). *Developing primary mathematics teaching. Reflecting on practice with the knowledge quartet.* London: Sage publications.

Rubin, A., Bruce B., & Tenney, Y. (1990). Learning about sampling: Trouble at the core of Statistics. *Proceedings of the ICOTS 3* (pp. 314–320). http://iase-web.org/

Sanchez, E., Silva, C., & Coutinho, C. (2011). Teachers' understanding of variation. In C. Batanero, G. Burrill & C. Reading (Eds.) *Teaching statistics in school mathematics – Challenges for teaching and teacher education. A joint ICMI/IASE study: The 18th ICMI Study* (pp. 211–222). New York, NY: Springer.

Savenye, W., & Robinson, R. (2005). Using qualitative research methods in higher education. *Journal of Computing in Higher Education*, *16*(2), 65–95.

Scheaffer, R. (2000). Statistics for a new century. In M. J. Burke & F. R. Curcio (Eds.), *Learning mathematics for a new century* (pp. 158–173). Reston, VA: NCTM.

Scheaffer, R. (2006). Statistics and Mathematics: On making a happy marriage. In G. Burrill (Ed.) *Thinking and reasoning with data and chance: Sixty-eighth yearbook of the NCTM* (pp. 309-322). Reston, VA: NCTM.

Schön, D. (1983). *The reflective practitioner: How professionals think in action*. Aldershot Hants: Avebury.

Seidman, I. (2006). *Interviewing as qualitative research – A guide for researchers in education and the social sciences*. New York: Teachers College Press.

Sharma, S. (2005). High school students' understanding of samples and sampling variability: Implications for teaching and research. In D. Clarkson, A. Downton, D. Gronn, M. Horne, A. McDonough, R. Pierce & A. Roche (Eds.), *Proceedings of the Twenty-eighth Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 517–524). Melbourne: MERGA. http://www.merga.net.au/documents/RP622004.pdf

Shaughnessy, J.M. (2006). Research on students' understanding of some big concepts. In G. Burrill (Ed.) *Thinking and reasoning with data and chance: Sixty-eighth yearbook of the NCTM* (pp. 77-98). Reston, VA: NCTM.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957–1000). Greenwich, CT: Information Age Publishing and NCTM.

Shaughnessy, J. M., & Chance, B (2005). *Statistical questions from the classroom*. Reston, VA: NCTM.

Shaughnessy, J. M., Chance, B., & Kranendonk, H. (2009). *Focus in high school mathematics. Reasoning and sense making*. Reston, VA: NCTM.

Shulman, L. (1986) Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.

Simon, M. (2006). Key developmental understandings in Mathematics: A direction for investigating and establishing learning goals. *Mathematical Thinking and Learning*, *8*(4), 359–371.

Silverman, J., & Thompson, P. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal of Mathematics Teacher Education*, *11*, 499–511.

Snee, R. (1993). What's missing in statistical education? *The American Statistician*, *47*(2), 149–154.

Sorto, A., White, A., & Lesser, L. (2011). Understanding student attempts to find a line of fit. *Teaching Statistics*, *33*(2), 49–52.

Sousa, M. O. (2002). *Investigações estatísticas no 2.º ciclo do ensino básico*. Dissertação de mestrado, DEFC, Universidade de Lisboa. Lisboa: APM.

Sowder, J. T. (2007). The mathematical education and development of teachers. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. A Project of the NCTM (vol I, pp. 157–223). Charlotte: Information Age Publishing.

Speer, N., King, K., & Howell, H. (2015). Definitions of mathematical knowledge for teaching: using these constructs in research on secondary and college mathematics teachers. *Journal Mathematical Teacher Education*, *18*, 105–122.

Stake, R. E. (1994). Case studies. In N. Denzin, & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236–247). London: Sage.

Stake, R. E. (2007). *A arte da investigação com estudos de caso*. Lisboa: Fundação Calouste Gulbenkian.

Torok, R., & Watson, J. (2000). Development of the concept of statistical variation: An exploratory study. *Mathematics Education Research Journal*, *12*(2), 147–169.

Turner, F. (2008). Growth in teacher knowledge: Individual reflection and community participation. In O. Figueras, J. Cortina, S. Alatorre, J. Rojano, & A. Sepúlveda (Eds.), *Proceedings of the 32nd conference of the International Group for the Psychology of Mathematics Education* (v. 4, pp. 356–360). Mexico: University of Saint Nicholas of Hidalgo.

Turner, F., & Rowland, T. (2011). The knowledge quartet as an organizing framework for developing and deepening teachers' mathematics knowledge. In T. Rowland & K. Ruthven (Eds.), *Mathematics knowledge in teaching* (pp. 195–212). Dordrecht: Springer.

Utts, J. (2003). What educated citizens should know about statistics and probability. *The American Statistician, 57*(2), 74–79.

Watson, J. (2001). Longitudinal development of inferential reasoning by school students. *Educational Studies in Mathematics, 47*, 337–372.

Watson, J. (2006). *Statistical literacy at school: Growth and goals*. New Jersey, Lawrence Erlbaum Associates Publishers.

Watson, A., & Harel, G. (2013). The role of teachers' knowledge of functions in their teaching: a conceptual approach with illustrations from two cases. *Mathematics and Technology Education, 13*(2), 154–168.

Watson, J., & Callingham, R. (2013). PCK and average. In V. Steinle, L. Ball & C. Bardini (Eds.), *Mathematics education: Yesterday, today and tomorrow.*

Proceedings of the 36th annual conference of the Mathematics Education Research group of Australasia (pp. 642–649). Melbourne, VIC: MERGA.

Watson, J., & Shaughnessy, J.M. (2004). Proportional reasoning: Lessons from research in data and chance. *Mathematics Teaching in the Middle School*, *10*(2), 104–109.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223–265.

Yin, R. E. (1989). *Case study research: Design and methods*. Newbury Park: Sage Publications.

Yin, R. E. (2003). *Case study research – design and methods*. Thousand Oaks: Sage Publications.

ANEXOS

ANEXO 1

Guiões das entrevistas

ENTREVISTA INICIAL (E1)

1. Percurso e contexto profissional

- Interesse pela Matemática e por ser professor.
- Informação sobre a formação inicial, tempo de serviço, escolas onde esteve colocada e níveis de ensino lecionados.
- Perspetiva sobre a Estatística enquanto aluna do ensino secundário e enquanto aluna do ensino superior. Conteúdos estudados. Dificuldades.
- Participação em projetos, clubs, encontros e formações. Contribuições dessas participações para a prática letiva.
- Conhecimentos estatísticos adquiridos e a sua origem.
- Caraterização geral do contexto profissional e da turma de 10.º ano, o trabalho realizado com o grupo disciplinar e relação com colegas e alunos.

2. Perspetiva geral sobre o conhecimento didático do professor em Estatística

- Planificação do tema no ensino secundário. Recursos.
- Principais estratégias e metodologias de ensino (em que diferem relativamente a outros temas da matemática). Das primeiras experiências às mais recentes.
- Conteúdos comuns na área da Estatística entre o programa de Matemática do 3.º ciclo e o programa de Matemática A do ensino secundário.
- Leitura ao tema da Estatística do Programa de Matemática A do ensino secundário. Tópicos estatísticos do programa de Matemática A do ensino secundário. A calculadora gráfica.
- Caracterização das tarefas propostas nas aulas. Investigações estatísticas. As quatro fases de um processo investigativo.
- Apreciação global relativamente às capacidades desenvolvidas nas aulas a partir do tipo de tarefa proposto.
- Estratégias desenvolvidas pelos alunos e dificuldades enfrentadas na resolução das tarefas. Atuação da professora no sentido de minimizar essas dificuldades.
- Utilização da tecnologia.

- Estatística é um tema curricular que devia ser mais desenvolvido?
- Perspetivas que predominam entre os professores sobre o ensino e a aprendizagem da Estatística.
- Análise e reflexão da professora sobre tarefas trazidas pela investigadora. Via-se a propor alguma delas? Desafios e dificuldades no ensino. Dificuldades na aprendizagem.
- Perspetiva da professora sobre a ideia com que os alunos ficam da Estatística após lecionação do tema.

GUIÃO DAS ENTREVISTAS PRÉ-AULA

- Conteúdos a abordar.
- Objetivos gerais da aula.
- Metodologia prevista, incluindo opções tomadas relativamente à seleção de tarefas.
- Dificuldades esperadas.

GUIÃO DAS ENTREVISTAS PÓS-AULA

- Aspetos que corresponderam à espectativa e os que se revelaram mais complexos.
- Interpretação da professora relativamente aos aspetos mais marcantes que ocorreram na aula.

ENTREVISTA 2 (E2)

Conhecimento didático da professora em Estatística

- Em cada tópico estatístico solicitou-se uma apreciação ao modo como alguns conceitos e representações foram compreendidos pelos alunos, às capacidades que os alunos deviam desenvolver, ao papel reservado à calculadora gráfica nas aulas. Adicionalmente, pediu-se à professora uma análise e reflexão sobre:

  - Preparação do tópico em termos de sequenciação de conteúdos, escolha e proveniência das tarefas (manual adotado, outros manuais, sites da internet, criadas ou adaptadas pela professora, etc). Características das tarefas propostas. Desafios e dificuldades enfrentados nessa preparação

  - Objetivo e potencialidade de cada tarefa no ensino. Expectativa. Dificuldades e desafios enfrentadas na implementação das tarefas. Reação geral dos alunos. Estratégias usadas pelos alunos. Dificuldades e erros encontradas na aprendizagem. Nível de conhecimento atingido pelos alunos.

  - Produções dos alunos realizadas nas aulas (excertos de gravações de aulas) e nos trabalhos de casa. Expectativa. Estratégias usadas pelos alunos. Erros e dificuldades encontrados. Identificação de aspetos a melhorar.

ENTREVISTA 3 (E3)

Conhecimento didático da professora em Estatística

- Em cada tópico estatístico solicitou-se mais alguns pedidos de apreciação relativamente a capacidades a desenvolver ao longo de cada tópico, ao modo como alguns conceitos e representações emergiram nas aulas e ainda quanto a entendimentos que os alunos desenvolveram. Também se questionou a professora sobre a sua compreensão relativamente a algumas noções estatísticas e suas propriedades, sobretudo a partir de situações que envolviam respostas de alunos (através das situações consideradas no anexo 2 e das tarefas propostas na aula). Adicionalmente, pediu-se à professora uma análise e reflexão sobre:

    - Alguns excertos de gravação de aulas centrados na exposição de noções estatísticas, na utilização da calculadora gráfica e nas intervenções de alunos quando lidavam com tarefas propostas (orais e/ou escritas). Expectativa. Aspetos mais bem conseguidos e os menos bem conseguidos nesses momentos e porquê? Estratégias usadas pelos alunos. Dificuldades encontradas. Identificação de aspetos a melhorar.
    - Produções escritas dos alunos realizadas nos trabalhos de casa. Expectativa. Estratégias usadas. Dificuldades e erros encontradas. Identificação de aspetos a melhorar.

- Também solicitou-se à professora a sua perspetiva sobre:

    - Ensino da Estatística na disciplina de Matemática A no 10.º ano agora que terminou a lecionação do tema.
    - Aprendizagem da Estatística na disciplina de Matemática A no 10.º ano agora que terminou a lecionação do tema.
    - Eventual influência da sua participação na ação de formação de Estatística para a sua prática (indicação de aspetos específicos).

- Eventuais aspetos a reconsiderar para uma próxima vez que lecione o tema de Estatística na disciplina de Matemática A.
- A sua participação neste estudo, um balanço final.

ANEXO 2

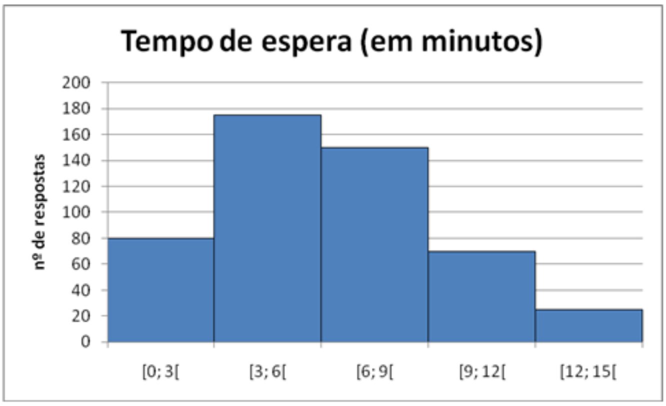
Situações apresentadas nas entrevistas

Tarefa *Média sobre média* (tarefa selecionada pela investigadora de um dos manuais adotados pelas professoras do estudo e utilizada na E1)

Um professor obteve, no primeiro teste de um período letivo, as médias 10,3; 15,2 e 13,4 em cada uma das suas três turmas do 10º ano, que têm, respetivamente, 32, 25 e 17 alunos. Determina a classificação média conjunta.

Tarefa *Tempo de espera pelo autocarro* (tarefa adaptada pela investigadora de um dos manuais adotados e utilizada na E1)

Numa cidade, em algumas paragens de autocarro, várias pessoas responderam à pergunta: "Quantos minutos espera habitualmente pelo autocarro?" As respostas obtidas foram sumariadas no histograma abaixo.



a) Quantas pessoas responderam à questão?

b) Constrói uma tabela de frequências relativas (simples e acumuladas)

c) O tempo de espera é considerado aceitável se for inferior a 6 minutos. Será a média deste conjunto de dados aceitável?

d) Calcule o desvio-padrão desta distribuição e interprete o resultado.

Tarefa *Como está a evoluir a nossa população*? (tarefa criada pela investigadora com dados demográficos portugueses de 2001 e de 2011, retirados do site INE e utilizada na E1)

Na tabela está registada a distribuição da população portuguesa por três grupos etários em 2001 e em 2011 (de acordo com os dados do *INE*). Faça um estudo à evolução da população portuguesa apoiando-se nas orientações seguintes e escreva as suas conclusões, interpretando a análise efetuada e relacionando-a com a questão de interesse - Como está a evoluir a nossa população?

População portuguesa (milhares)	2001	2011
***Pop. jovem* (0-14 anos)**	1 679 492	1 572 900
***Pop. em idade ativa* (15-65 anos)**	6 992 760	6 961 852
***Pop. idosa* (65 e mais anos)**	1 722 417	2 007 646

Orientações:
- Procure padrões/relações nos dados;
- Analise cada uma das distribuições;
-Compare as distribuições;
- Analise a variabilidade dos dados;
-Use representações gráficas diversas e medidas estatísticas na sua análise, tendo o cuidado de interpretar todos os resultados;
- Faça conjeturas/inferências acerca do contexto dos dados.

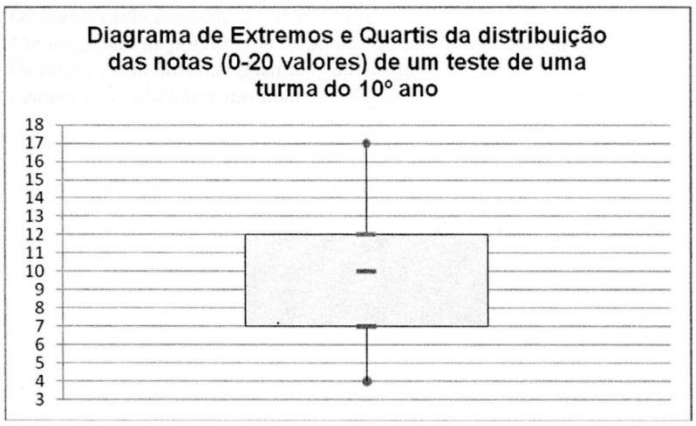
Situação *Desvio-padrão elevado* (criada pela investigadora e utilizada na E3)

O que poderia significar um desvio-padrão elevado para um conjunto de dados?

1. Os dados estão bastante dispersos entre si;
2. Um diagrama de pontos com os pontos a variar bastante entre si;
3. Os dados bastante afastados da média;
4. Pequena variabilidade nos dados com a existência de um ou mais *outliers*.

Situação *Respostas hipotéticas sobre um diagrama de extremos e quartis* (criada pela investigadora e utilizada na E3)

O professor de matemática mostrou na turma com 24 alunos o seguinte diagrama de extremos e quartis relativamente à distribuição das notas da turma no último teste realizado. E pediu aos alunos para fazerem comentários ao diagrama.



1. *Esta turma é razoável pois tem exactamente 50% de notas positivas.*
2. *Há 25% de notas a variar entre 4 e 7 valores.*
3. *Há pelo menos 50% dos alunos que tiraram notas iguais ou superiores a 10.*
4. *A amplitude é de 13 valores, existe muita dispersão de dados.*
5. *Há uma grande concentração de resultados entre 10 e 12 valores.*
6. *Há uma maior dispersão de notas superiores ou iguais ao terceiro quartil.*

ANEXO 3

Tarefas propostas e situações apresentadas na aula de Estela

T1: Tarefa *Satisfação dos clientes* (selecionada de um manual e integrada numa ficha de trabalho)

A empresa Preguicite Aguda, S.A. pretendia elaborar um estudo sobre a satisfação dos seus clientes relativamente aos serviços prestados. Para tal, realizou um inquérito a 100 clientes escolhidos aleatoriamente.

a) Indique a população, a amostra e a unidade estatística.

b) Diga se a amostra escolhida é ou não representativa da população em questão. Quais são as condições para que uma amostra seja representativa?

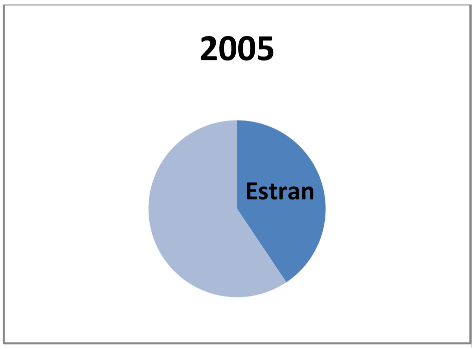
c) O estudo efectuado trata-se de um censo ou de uma sondagem?
Justifique.

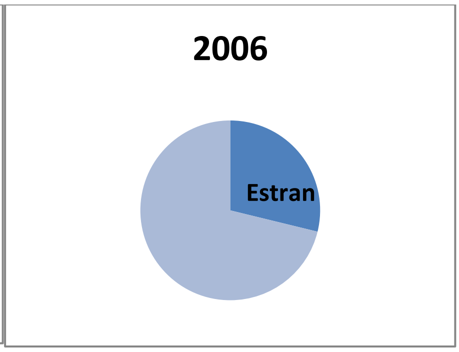
d) Sabendo que os clientes interrogados podiam optar pelas respostas "totalmente insatisfeito", "não satisfeito", "satisfeito" e "plenamente satisfeito", indique e classifique a variável estatística em estudo.

T2: Tarefa *Visitas a um museu* (selecionada do manual adotado)

Num relatório sobre as visitas a um museu em dois anos consecutivos, 2005 e 2006, apresentavam-se os dados e gráficos seguintes:

Tabela	2005	2006
Visitantes estrangeiros	9563	12078
Total de visitantes	23551	41915





Este relatório foi devolvido para retificação por motivo de incoerência, acompanhado do seguinte comentário: *A tabela indica que houve um aumento de visitantes estrangeiros, enquanto os gráficos demonstram uma redução.* Interprete o comentário e dê a sua opinião sobre a situação descrita e sobre o procedimento a adotar em relação ao relatório.

T3: Tarefa *Número de cães examinados* (selecionada de um manual e integrada numa ficha de trabalho)

O número de cães que foram examinados, em cada um dos dias ao longo de 25 dias, numa Clínica Veterinária foram os seguintes:

25	20	30	15	20	40	35	40	30	25	20	25	30
15	35	20	40	25	15	20	25	15	25	30	35	

a) Indique e classifique a variável estatística em estudo.
b) Construa uma tabela de frequências, que contenha as frequências absolutas, as frequências relativas, as frequências absolutas acumuladas e as frequências relativas acumuladas, com base nos dados fornecidos.
c) Qual a percentagem de dias em que foram examinados 25 cães? (Sugestão: calcule a frequência relativa em percentagem correspondente).
d) Construa um gráfico de barras que represente a distribuição dos dados anteriores, considerando as frequências absolutas.

T4: Tarefa *Classificações a Matemática das turmas A e B* (selecionada de um manual e integrada numa ficha de trabalho)

As classificações dos testes de Matemática da turma A do 10º ano, na escala de 0 a 20, foram as seguintes:

Turma A				
12	13	15	12	11
9	10	16	17	13
7	9	10	19	13
15	6	14	13	9
11	18	11	8	14

As classificações da turma B distribuem-se de acordo com o seguinte diagrama de caule – e – folhas:

Turma B

0	7 8 9 9 9
1	0 0 0 1 2 3 4 4 5 6 7 8 8
2	0 0

a) Apresenta a distribuição das classificações da turma A num diagrama de caule – e – folhas
b) Qual das turmas apresenta maior percentagem de negativas?
c) Atribui-se “Muito Bom” a classificações não inferiores a 18. Quantos se registaram em cada turma?

T5: Tarefa *Idades de jovens num casting* (selecionada do manual adotado)

As idades de 30 jovens actores presentes num *casting* para uma comédia musical foram as seguintes:

21	19	22	19	18	20	23	19	19	20
19	20	21	22	21	20	22	20	21	30
21	19	21	21	19	19	20	19	19	19

Construa uma tabela de frequências absolutas e relativas e represente, recorrendo a um diagrama de caule-e-folhas, este conjunto de dados.

Represente graficamente a função cumulativa.

T6: Tarefa *Ração gasta num dia* (selecionada de um manual e integrada numa ficha de trabalho)

O Sr. A. Galo é diretor de um aviário, e registou o número de quilogramas de ração que foram gastos durante os primeiros 15 dias do mês de Abril; os dados obtidos foram os seguintes:

30,3	25,8	20,3	21,9	29,2	35,1	27,3	19,7	32,6
37,5	40,3	29,2	32,4	33,7	38,4			

a) Identifique e classifique a variável estatística em estudo.
b) Sabendo que se consideram 5 classes com a mesma amplitude e que as primeiras 4 são [19,5;23,7[, [23,7;27,9[, [27,9;32,1[e [32,1;36,3[, determine a última classe.
c) Construa a tabela de frequências absolutas, simples e acumuladas .
d) Qual o número de dias em que o número de quilogramas de ração gastos foi inferior a 32,1?
e) Determina a média do número de quilogramas gastos durante os primeiros 15 dias de Abril.

T7: Tarefa *Número de hóspedes de uma estância balnear* (selecionada de um manual e integrada numa ficha de trabalho)

A distribuição seguinte representa o número de hóspedes de uma estância balnear, em determinada semana: 90, 80, 80, 100, 110, 240, 180.

a) Determina a média, a moda e a mediana do número de hóspedes da estância balnear nessa semana.
b) Calcule os quartis e apresente o diagrama de extremos e quartis.
c) Confirme os valores e o gráfico encontrados, na calculadora gráfica.

T8: Tarefa *Temperaturas máximas de Lisboa e de Madrid* (selecionada de um manual e integrada numa ficha de trabalho)

A tabela que a seguir se apresenta refere-se às temperaturas máximas atingidas nas cidades de Lisboa e Madrid, durante a semana de 8 a 14 de Maio.

Lisboa	27	26	22	23	24	25	25
Madrid	28	28	22	21	21	24	23

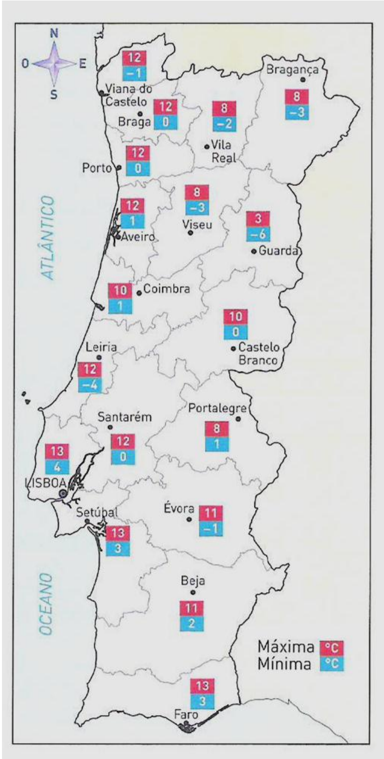
a) Determina a média, moda e mediana para cada uma das distribuições.
b) Calcula o desvio padrão para cada um dos casos e interpreta os valores obtidos relativamente à situação indicada.

T9: Tarefa *Temperaturas máximas e mínimas de 18 cidades* (selecionada de um manual e integrada numa ficha de trabalho)

Num dia de Inverno, as temperaturas mínimas e máximas registadas nas dezoito cidades, capitais de distrito de Portugal Continental, estão indicadas no mapa da figura.

a) Com os dados referentes às dezoito cidades referidas no mapa preenche uma tabela do tipo da que é indicada a seguir.

Cidade	Temp. máx. T	Temp. mín. t	Ampl. térmica T - t
Viana	12	-1	13
…	…	…	…
Faro	13	3	10

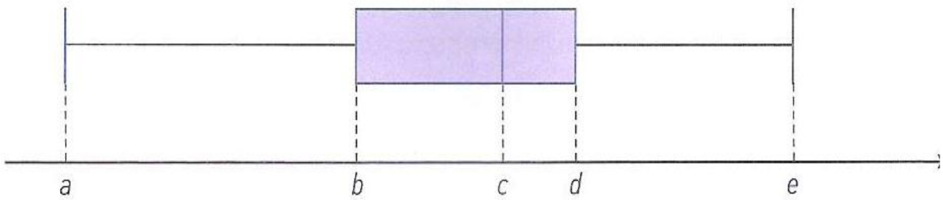


b) Em que cidade se registou menor amplitude térmica? E maior?

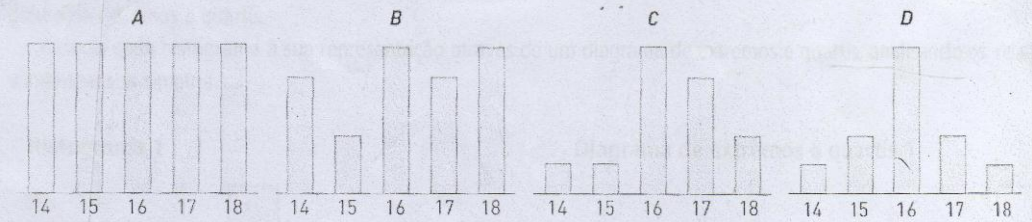
c) Calcula a média aritmética das amplitudes térmicas, apresentando o resultado arredondado às décimas.

d) Considera a distribuição das temperaturas máximas. Determina os quartis e constrói o diagrama de extremos e quartis.

e) Considera a distribuição das temperaturas mínimas. Em relação a esta distribuição, o diagrama de extremos e quartis a seguir apresentado encontra-se incompleto. Indica os valores de ***a***, ***b***, ***c***, ***d*** e ***e***.



T10: Tarefa *Distribuições das idades das equipas de trabalho* (selecionada de um manual)



Considera os seguintes pares de valores $(\bar{x}\ ,\ s)$ relativos às distribuições das idades dos elementos dos grupos de trabalho.

I – $(\bar{x}\ ,\ s) = (16\ ;\ 1{,}05)$ II – $(\bar{x}\ ,\ s) = (16{,}3\ ;\ 1{,}14)$

III – $(\bar{x}\ ,\ s) = (16\ ;\ 1{,}41)$ IV – $(\bar{x}\ ,\ s) = (16{,}3\ ;\ 1{,}41)$

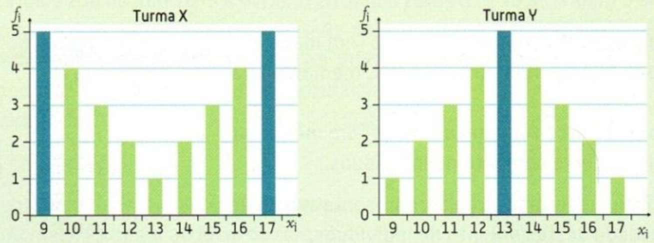
A correspondência entre pares de valores $(\bar{x}\ ,\ s)$ (média e desvio-padrão) e os gráficos *A*, *B*, *C* e *D* é:

(A) I – *D*, II – *C*, III – *A*, IV – *B*; (B) I – *A*, II – *C*, III – *D*, IV – *B*;

(C) I – *D*, II – *B*, III – *A*, IV – *C*; (D) I – *C*, II – *D*, III – *A*, IV – *B*.

T11: Tarefa *Aproveitamento das turmas X e Y (*selecionada do manual adotado)

Os gráficos representam as classificações obtidas pelos alunos de duas turmas para uma mesma disciplina.



Qual a classificação média de cada turma?

Sem fazer cálculos, preveja qual a turma que tem maior desvio-padrão das classificações. Justifique.

Calcule o desvio-padrão, a mediana e os quartis de ambas as turmas.

Compare o aproveitamento das duas turmas servindo-se das medidas estatísticas calculadas.

T12: Tarefa Relatório sobre os resultados do teste intermédio (criada pela professora e integrada numa ficha de trabalho)

Considera a distribuição das classificações obtidas no TI do 10ºA de uma escola secundária. Recorrendo aos teus conhecimentos de estatística elabora um relatório acompanhado de gráficos e valores estatísticos que consideres relevantes para o estudo.

Classif.	N.º Aluno	%
0	0	0
1	0	0
2	0	0
3	0	0
4	0	0
5	2	8
6	4	16
7	0	0
8	1	4
9	4	16
10	2	8
11	2	8
12	5	20
13	2	8
14	2	8
15	0	0
16	0	0
17	0	0
18	1	4
19	0	0
20	0	0

T13: Tarefa *Propriedades da média e do desvio-padrão* (adaptada de uma tarefa de um manual e integrada numa ficha de trabalho)

Parte 1:

Numa empresa os salários em 2005 eram os constantes na tabela seguinte:

Salário (euros)	Freq. abs.
600	20
700	50
800	15
1000	9
1500	3
2000	2
4000	1

a) Determina a média, a moda, a mediana, os quartis e o desvio-padrão da distribuição;
b) Supondo que houve um aumento de 5% nos salários em 2006, quais serão os valores dos salários em 2006?
c) Determina a média, a moda, a mediana, os quartis e o desvio-padrão da distribuição em 2006 com um aumento de 5%. Que conclusões tiras, comparando com os resultados de a) ?
d) Supondo que houve um aumento de 50 euros nos salários em 2006, quais serão os valores dos salários em 2006?
e) Determina a média, a moda, a mediana, os quartis e o desvio-padrão da distribuição em 2006 com um aumento de 50 euros. Que conclusões tiras, comparando com os resultados de a)?

Parte 2

Propriedades da média

1) Ao somar a todos os dados da distribuição de média $\bar{x}$ a mesma quantidade K, então a nova distribuição de dados terá média $\bar{y} =$______________________________

2) Ao multiplicar todos os dados da distribuição de média $\bar{x}$ a mesma quantidade K, então a nova distribuição de dados terá média $\bar{y} =$____________________________

Propriedades do desvio-padrão

1) Ao somar a todos os dados da distribuição de desvio-padrão σ a mesma quantidade K, então a nova distribuição de dados terá desvio-padrão________________

2) Ao multiplicar todos os dados da distribuição de desvio-padrão σ a mesma quantidade K, então a nova distribuição de dados terá desvio-padrão________________

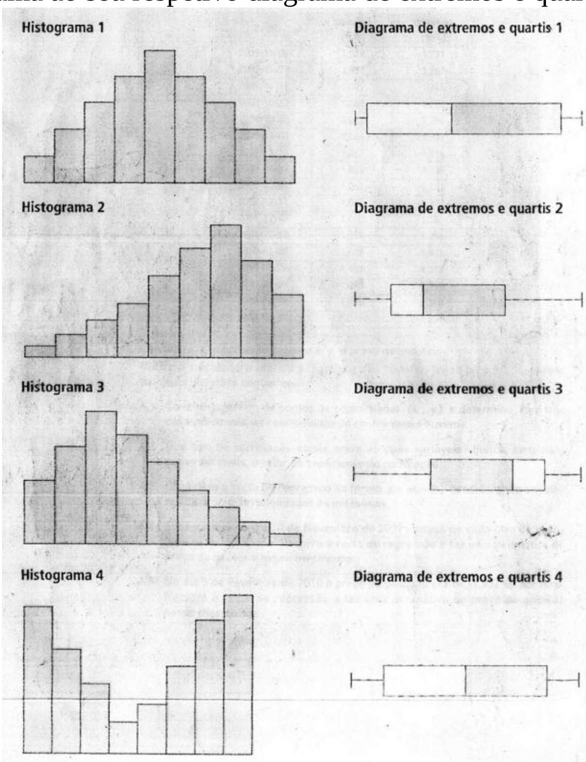
T14: Tarefa *Alteração dos dados de uma distribuição* (criada pela professora e integrada numa ficha de trabalho)

- Supõe que se adicionou 20 a cada um dos valores de uma dada distribuição. O que acontece à média, moda, mediana, amplitude total, amplitude interquartis e ao desvio padrão?

- Supõe que se multiplicou por 20 cada um dos valores de uma dada distribuição. O que acontece à média, moda, mediana, amplitude total, amplitude interquartis e ao desvio-padrão?

T15: Tarefa *Associa entre histogramas e diagramas de extremos e quartis* (selecionada de um manual)

Associe cada histograma ao seu respetivo diagrama de extremos e quartis.



T16: Tarefa *A equipa de basquetebol do Porto* (adaptada de um manual e integrada numa ficha de trabalho)

Na tabela estão indicados a idade (I), a altura (A), e as médias por jogo, dos minutos em campo (M) e dos pontos marcados (P), da equipa de Basquetebol do Porto na presente época (2000/2001) até à 13ª jornada segundo dados recolhidos no site da Infordesporto: http://www.infordesporto.pt

Jogador	Idade	Altura	Min./Jogo	Pontos/Jogo
Anthony Blackely	35	2.04	30	12.1
Elvis Évora	22	2.05	24.4	9.4
José Pedrera	28	2.02	22.4	9.6
Nuno Marçal	25	2.05	23.8	15
Nuno Perdigão	22	1.92	21.8	8.1
Nuno Quidiongo	24	1.80	5.9	4
Kevin Vulin	25	2.04	31.2	15.4
Paulo Cunha	20	2.00	3.5	1
Rui Santos	30	1.88	30.1	8.7
João Rocha	25	2.00	10.9	4.1
Raúl Santos	31	2.03	3.8	1.4
Bob Harsad	31	1.98	28.8	14.9

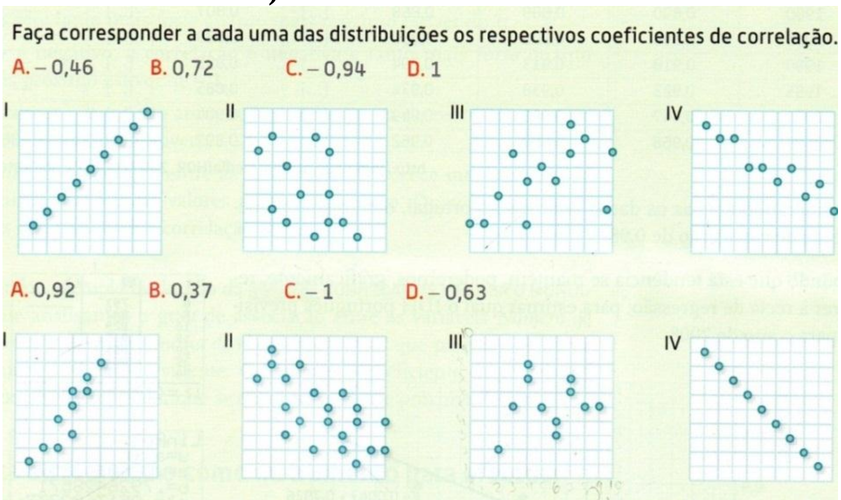
1- Será que existe alguma relação entre a altura do jogador e os minutos que este está em campo? Dito de outra forma, será que os jogadores mais altos são solicitados mais vezes a jogar?

2- E entre a idade e os pontos que marca? Será que os jogadores mais novos marcam mais pontos?

3- E quanto à eficácia do jogo, será que existe relação entre os minutos de jogo e os pontos obtidos?

Com a ajuda da calculadora representa os diagramas de dispersão e tenta dar resposta às questões anteriores.

T17: Tarefa *Associação entre nuvens de pontos e coeficientes de correlação* (selecionada do manual adotado)

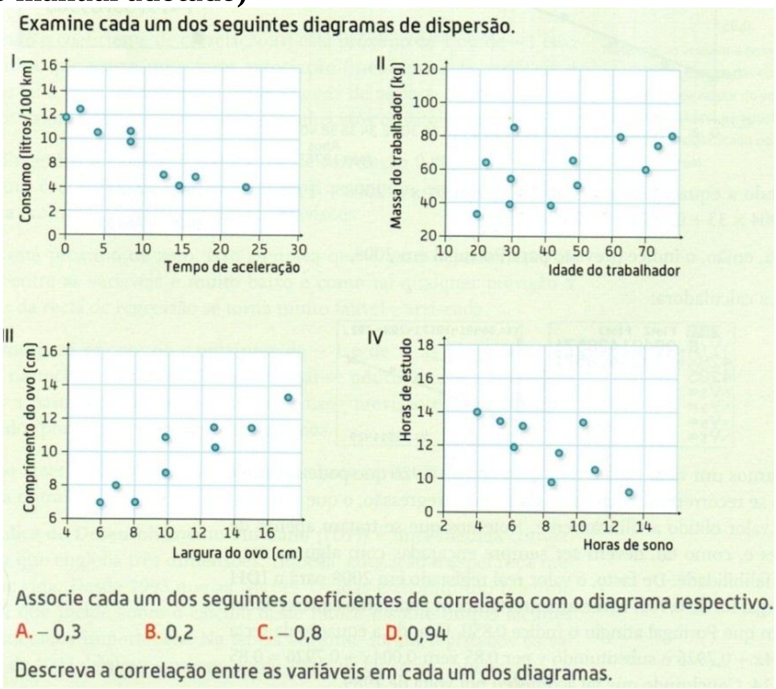


Faça corresponder a cada uma das distribuições os respectivos coeficientes de correlação.

A. – 0,46 B. 0,72 C. – 0,94 D. 1

A. 0,92 B. 0,37 C. – 1 D. – 0,63

T18: Tarefa *Associação entre diagramas de dispersão a coeficientes de correlação* (selecionada do manual adotado)



Examine cada um dos seguintes diagramas de dispersão.

Associe cada um dos seguintes coeficientes de correlação com o diagrama respectivo.

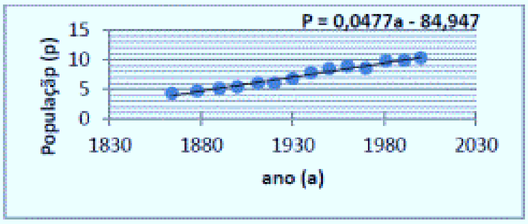
A. – 0,3 B. 0,2 C. – 0,8 D. 0,94

Descreva a correlação entre as variáveis em cada um dos diagramas.

T19: Tarefa *População residente em Portugal* (selecionada do manual adotado)

Na tabela estão alguns dados sobre a população residente em Portugal desde 1864 até 2000. O diagrama de dispersão relativo a estes dados assim como a respectiva reta de regressão e a sua equação estão apresentados na figura abaixo.

Ano	População	Ano	População
1864	4,3	1940	7,8
1878	4,7	1950	8,5
1890	5,1	1960	8,9
1900	5,4	1970	8,6
1911	6,0	1981	9,8
1920	6,0	1991	9,9
1930	6,8	2000	10,3



Explique por que razão o modelo linear apresentado não pode ser adequado para: (a) estimar o número aproximado de habitantes, em Portugal, há uns séculos e (b) prever a evolução da população portuguesa, a muito longo prazo (relacione uma tal previsão com os recursos, alimentares e outros, necessariamente limitados).

T20: Tarefa *Estudo com carateres qualitativos e quantitativos* (criada pela professora).

Completa:

Nome do aluno	*Cor dos olhos*	*Idade (anos)*	*Nível a Mat. - nota final de 9.º ano*	*Altura (metros)*	*Peso (Kgs)*

T21: Tarefa *Massa corporal feminina* (selecionada de um manual)

Relativamente a um grupo de 15 mulheres portuguesas, fizeram-se os registos da altura e massa corporal de cada uma delas, tendo-se obtido os seguintes valores:

Altura	157	160	165	159	168	172	170	165	166	163	169	159	169	168	168
Massa	52	61	62	60	67	70	69	66	65	61	58	72	72	68	66

a) Investiga a existência de correlação entre as variáveis em estudo.
b) Identifica o centro da nuvem de pontos.
c) Determina a reta de regressão.
d) Qual a massa corporal que se prevê para uma mulher com 175 cm de altura?

Situação *Mediana na tabela* (selecionada do e-escola)

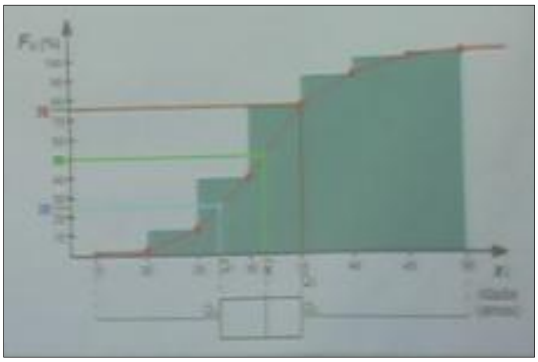
Perguntou-se a um grupo de 115 mulheres

o número que cada uma delas calçava…

Xi	***fi***	***Fi***
35	***2***	***2***
36	***10***	***12***
37	***26***	***38***
38	***30***	***68***
39	***29***	***97***
40	***17***	***114***
41	***1***	***115***

Situação *Quartis em gráficos* (selecionado do e-escola)

Visualização dos quartis num gráfico que inclui polígono/histograma de frequências - cujos dados correspondem às idades de um grupo de pessoas (fornecidos em classes).



ANEXO 4

Tarefas propostas e situações apresentadas na aula de Lia

T1: Tarefa *Intenções de voto em 2006* (selecionada de um manual e incluída numa ficha de trabalho)

Lê a notícia e observa o gráfico da figura:

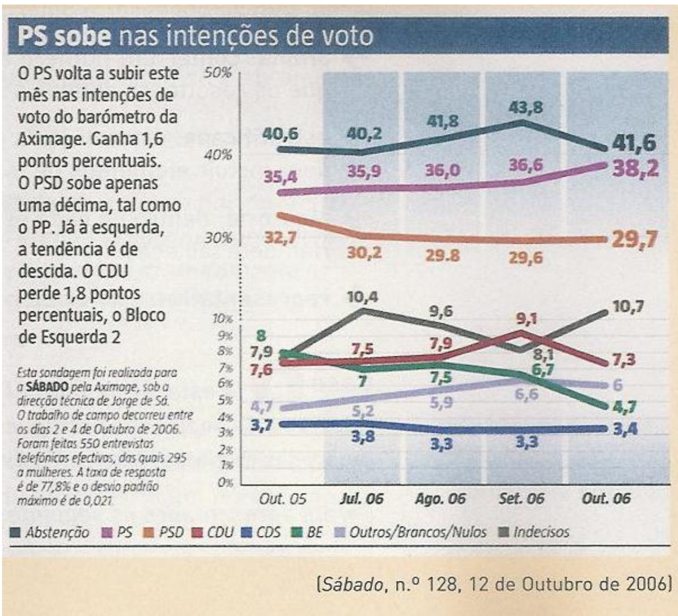
a) Este estudo é uma sondagem ou um recenseamento? Justifica.

b) De que forma foram inquiridos os elementos da amostra? E qual a sua dimensão?

c) Quantos homens foram entrevistados? Calcula a respetiva percentagem.

d) De Julho até Outubro de 2006, quantos pontos percentuais ganhou o PS? E quantos perdeu o PSD?

e) Quantas pessoas correspondem aos 77,8% da taxa de resposta?



T2: Tarefa *Alunos inscritos no ensino* (selecionada de um manual e incluída numa ficha de trabalho)

a) Indica em que países o ensino vocacionado representa uma percentagem superior a 70%.

b) Indica dois países que têm maior percentagem de estudantes no Ensino em Geral.

c) Em Portugal, dada a preocupação com as taxas de abandono escolar e com o número reduzido de alunos a frequentar Cursos Profissionais, o Ministério da Educação pretende conhecer o abandono escolar por parte dos alunos inscritos em cursos profissionais.

Para realizar um estudo, o ME poderá recorrer a uma sondagem ou a um recenseamento.

- Qual te parece ser o mais adequado?
- Analisa e regista vantagens e desvantagens na escolha de cada um dos tipos de estudo.

T3: Tarefa *Serão boas amostras?* (selecionada de um manual e incluída numa ficha de trabalho)

Veriica, justificando, se os exemplos seguintes se referem a boas amostras:

a) Usar utilizadores de um autocarro que pare junto a uma escola, para determinar a idade dos utilizadores dos transportes públicos.

b) Usar dez benfiquistas para prever o resultado do próximo jogo Porto-Benfica.

T4: Tarefa *Desportos praticados por 50 alunos de uma escola* (selecionada de um manual e incluída numa ficha de trabalho)

Numa escola secundária foram inquiridos 50 alunos sobre a sua prática desportiva. Registaram-se as seguintes respostas:

A B F O O B F N N N O N B B F F F F F F F F F F F

F F B A A B B N N N B A A N F N F A B A F N F F O

(F = Futebol; A = Andebol; B = Basquetebol; N = Natação; O = Outras)

Organize estes dados através numa tabela de frequências.

T5: Tarefa *Frequência com que os alunos de uma escola comem fast food* (selecionada de um manual e incluída numa ficha de trabalho)

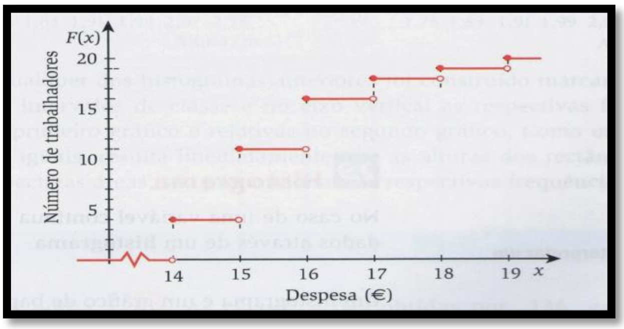
O conselho executivo de uma escola, preocupado com o excesso de peso de alguns alunos, resolveu fazer um estudo sobre o número de vezes por semana que os alunos comem *fast food*. Os resultados foram apresentados no seguinte gráfico de barras:

a) Regista numa tabela os dados representados no gráfico.

b) Incluí nessa tabela frequências absolutas e relativas acumuladas

T6: Tarefa *Despesas em transportes de 20 trabalhadores* (selecionada de um manual e incluída numa ficha de trabalho)

A 20 trabalhadores de uma empresa perguntou-se qual era a despesa que tinham, em euros, por semana, em transportes.



a) Define analiticamente a função cumulativa.

b) Quantos trabalhadores gastam, por semana, no máximo 18 euros?

c) Quantos trabalhadores gastam, por semana, exatamente 16 euros?

T7: Tarefa *Resultados do campeonato de atletismo nos 80 metros* (selecionada de um manual e incluída numa ficha de trabalho)

Depois de ordenados, os resultados deste campeonato foram:

80 m – Iniciados masculinos	10,45	10,48	10,57	10,59	10,67	10,75	10,88	10,89	10,90
	10,97	11,02	11,03	11,13	11,13	11,38	12,05	13,20	13,81
80 m – Iniciados femininos	11,51	11,54	12,18	12,18	12,30	12,45	12,65	12,96	12,98
	13,15	13,16	13,36	13,41	14,00	14,64			

Quantas classes se devem definir?

Existe uma fórmula que nos fornece uma indicação do número de classes a considerar:

O número de classes, k, a considerar numa amostra de dimensão n é o menor número inteiro tal que

$$2^k \geqslant n$$

a) Constrói a tabela de frequências absoluta simples (n_i) e relativa e, ainda, frequência absoluta acumulada (N_i).
b) Constrói o histograma e traça o respetivo polígono de frequências absolutas.
c) Constrói o histograma de frequências acumuladas e traça o respetivo polígono de frequências acumuladas.

T8: Tarefa *Golos marcados na liga de Sagres de futebol* (selecionada do manual adotado)

Dos jogos realizados na Liga Sagres de Futebol, em 2008-2009, seleccionou-se uma amostra constituída pelos jogos decorridos em duas jornadas consecutivas.

Pretende-se fazer o estudo estatístico sobre o número de golos marcados em cada jogo.

Os resultados dos jogos que fazem parte da amostra estão representados a seguir.

1.ª jornada

Académica	0 - 0	Leixões
Rio Ave	2 - 4	V. Guimarães
V. Setúbal	0 - 2	Sporting
E. Amadora	1 - 0	Naval
Marítimo	4 - 1	P. Ferreira
Nacional	2 - 4	FC Porto
Trofense	2 - 0	Benfica
Sp. Braga	2 - 0	Belenenses

2.ª jornada

V. Guimarães	0 - 0	E. Amadora
Sporting	2 - 0	Marítimo
Naval	2 - 1	Académica
Leixões	0 - 0	V. Setúbal
P. Ferreira	2 - 3	Nacional
FC Porto	0 - 0	Trofense
Benfica	1 - 0	Sp. Braga
Rio Ave	0 - 1	Belenenses

Qual é a dimensão da amostra escolhida?

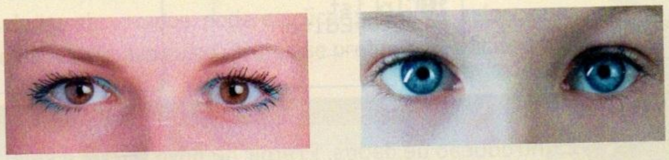
Identifica e classifica a variável em estudo.

Quais são os valores que assume a variável?

Constrói uma tabela de frequências absolutas e relativas.

T9: Tarefa *Cor dos olhos de 40 pessoas* (selecionada do manual adotado)

Num curso de fotografia inscreveram-se 40 jovens e fez-se o registo da cor dos olhos de cada um destes jovens.



A partir dos resultados obtidos foi construído o diagrama circular apresentado a seguir.

a) Dos jovens que frequentam o curso, quantos têm olhos verdes?

b) Faça uma tabela de frequências absolutas e relativas, para a variável cor dos olhos, indicando as frequências relativas em forma de fração irredutível.

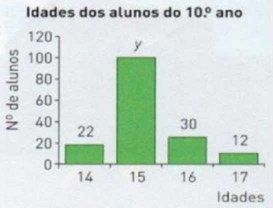
T10: Tarefa *Pictograma do número de irmãos* (selecionada do manual adotado)

N.º de irmãos	 = 4 alunos
0	
1	
2	
3	

a) Quantos alunos tem a turma?
b) Qual a percentagem de alunos da turma que têm mais de um irmão?

T11: Tarefa *Idade dos alunos de 10.º ano* (selecionada do manual adotado)

A distribuição das idades dos 160 alunos do 10.º ano de uma escola encontra-se representada no seguinte gráfico.



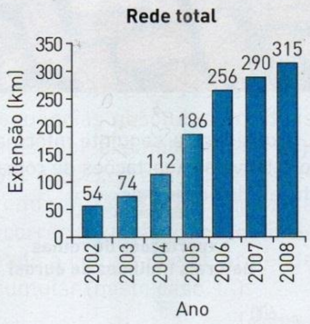
Determina o número de alunos com 15 anos de idade.

Copia para o teu caderno a seguinte tabela de frequências absolutas e relativas e completa-a.

Idade (anos)	Frequência absoluta n_i	Frequência relativa f_i
14	...	...
15	...	...
16	...	...
17	...	...

T12: Tarefa *Evolução da rede de saneamento* (selecionada do manual adotado)

Admite que a evolução da rede de saneamento existente num determinado concelho está representada no seguinte gráfico:



Em que ano se verificou a construção de maior quantidade de rede de saneamento?

Que quantidade de rede foi construída depois do final do ano 2005 ?

Elabora um quadro de frequências absolutas e relativas.

T13: Tarefa *Notas de Matemática das turmas A e B* (selecionada do manual adotado)

As classificações dos testes de Matemática da turma A do 10º ano, na escala de 0 a 20, foram as seguintes:

Turma A

12	13	15	12	11
9	10	16	17	13
7	9	10	19	13
15	6	14	13	9
11	18	11	8	14

As classificações da turma B distribuem-se de acordo com o seguinte diagrama de caule – e – folhas:

Turma B

0	7 8 9 9 9
1	0 0 0 1 2 3 4 4 5 6 7 8 8
2	0 0

a) Apresenta a distribuição das classificações da turma A num diagrama de caule – e – folhas.
b) Qual das turmas apresenta maior percentagem de negativas?
c) Atribui-se "Muito Bom" a classificações não inferiores a 18. Quantos se registaram em cada turma?

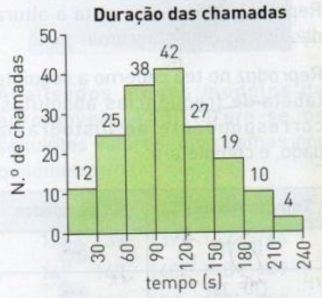
T14: Tarefa *Emissões de dióxido de carbono* (selecionada do manual adotado)

Analisados alguns modelos de automóveis, registaram-se os seguintes valores, em gramas por quilómetro:

128	130	132	134	134
138	139	139	142	144
147	147	148	148	149
149	150	153	155	158
158	160	162	163	165
165	167	167	168	169
170	171	172	172	172
173	173	174	174	175
175	192	193	193	194
196	196	197	199	199

Constrói uma tabela de frequências com os dados organizados em classes, aplicando a **regra empírica** referida no texto ao lado.

T15: Tarefa *Despesas em chamadas de um escritório* (Selecionada do manual adotado)

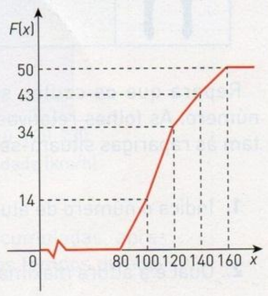


Constrói uma tabela de frequências correspondente ao gráfico apresentado. Apresenta as frequências relativas arredondadas às centésimas.

Qual é a percentagem de chamadas de duração inferior a 2 minutos?

T16: Tarefa *Áreas de vivendas* (Selecionada do manual)

Numa zona habitacional, a distribuição das áreas das vivendas, em m^2, é apresentada através do seguinte gráfico da função cumulativa:



Determina a percentagem de habitações com área inferior a $120\ m^2$.

Reproduz no teu caderno a seguinte tabela e, com base na informação dada, completa-a.

Área (m^2)	N.º de vivendas
[80 , 100[	...
[100 , 120[	...
[120 , 140[	...
[140 , 160[	...

T17: Tarefa *Qual a melhor turma, A ou B?* (tarefa criada pela professora e incluída numa ficha de trabalho)

As classificações, em pontos, de duas turmas de 10º ano num teste de avaliação, estão representadas na seguinte tabela:

Turma A	**10**	**20**	**30**	**31**	**41**	**45**	**45**	**53**	**76**	**76**	**90**	**90**	**90**
	97	**97**	**98**	**100**	**100**	**117**	**118**	**119**	**120**	**130**			
Turma B	**10**	**27**	**35**	**38**	**39**	**40**	**42**	**43**	**44**	**49**	**53**	**59**	**60**
	62	**68**	**69**	**70**	**76**	**77**	**94**	**190**	**190**	**198**	**199**		

a) Para cada uma das turmas A e B calcula: <u>Média; Moda; Mediana.</u>

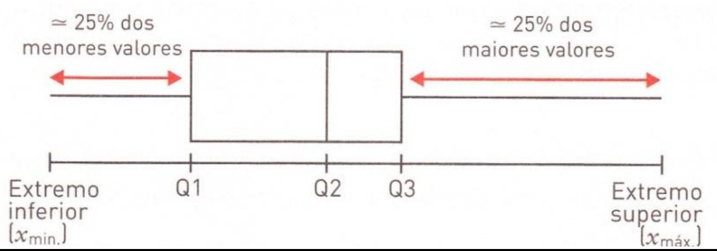
b) Para cada uma das turmas A e B calcula: <u>1ºQuartil; 3º Quartil.</u>

c) Para cada uma das turmas constrói um diagrama de extremos e quartis.

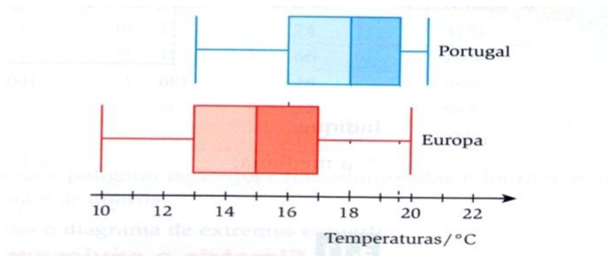
d) Analisa os resultados obtidos pelas duas turmas neste teste de avaliação e compara-os, tendo em conta todas as medidas de localização que determinaste.

NOTA: Os quartis são valores que dividem a distribuição em quatro partes iguais, cada uma delas com, aproximadamente, 25% dos dados. O 1º quartil representa-se por Q1; o 2º quartil representa-se por Q2 e coincide com a mediana; o 3º quartil representa-se por Q3. Um diagrama de extremos e quartis é uma representação gráfica que nos dá a informação do valor dos quartis e dos extremos inferior e superior da distribuição:

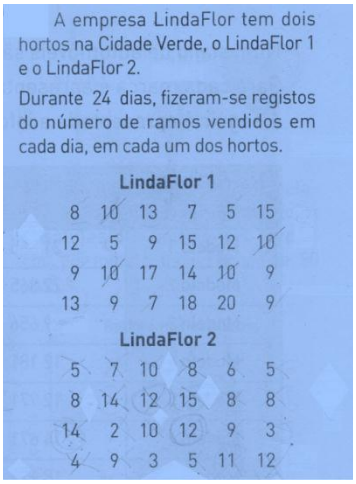
Estes gráficos permitem conhecer a forma como os dados se distribuem:



T18: Tarefa *Compara temperaturas máximas na Europa* (adaptada de um manual e incorporada numa ficha de trabalho) Considera os diagramas de extremos e quartis relativos a temperaturas máximas observadas durante um dia, nas 15 capitais europeias, excluindo Lisboa, e temperaturas máximas observadas nesse mesmo dia nas 18 capitais distritais de Portugal Continental. **Compara os diagramas e retira algumas conclusões.**



T19: Tarefa *Associa entre conjuntos de dados discretos (hortos) e diagramas de extremos e quartis* (selecionada do manual adoptado).


A empresa LindaFlor tem dois hortos na Cidade Verde, o LindaFlor 1 e o LindaFlor 2.

Durante 24 dias, fizeram-se registos do número de ramos vendidos em cada dia, em cada um dos hortos.

LindaFlor 1

8	10	13	7	5	15
12	5	9	15	12	10
9	10	17	14	10	9
13	9	7	18	20	9

LindaFlor 2

5	7	10	8	6	5
8	14	12	15	8	8
14	2	10	12	9	3
4	9	3	5	11	12

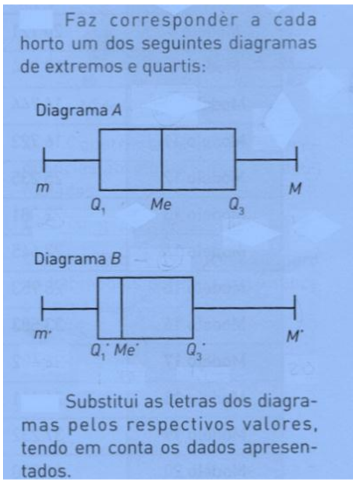

Faz corresponder a cada horto um dos seguintes diagramas de extremos e quartis:

Diagrama *A*

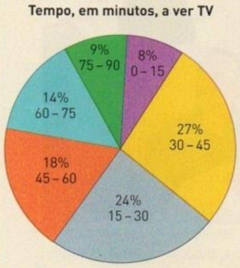
Diagrama *B*

Substitui as letras dos diagramas pelos respectivos valores, tendo em conta os dados apresentados.

T20: Tarefa *Inquérito a 500 alunos sobre televisão* (selecionada do manual adotado)

Numa escola, fez-se um inquérito a 500 alunos para se saber qual o tempo (em min) que dedicam diariamente a ver televisão.

Depois dos dados organizados, foi apresentado o seguinte diagrama circular com os resultados:



1. Dos alunos que responderam ao inquérito quantos gastam diariamente menos de 45 minutos a ver televisão?

2. Transcreve a tabela seguinte e completa-a.

Tempo (min)	Marca da classe x_i	n_i	N_i	f_i	F_i
0-15	...	...	...	...	...
15-30	...	...	...	...	...
30-45	...	...	...	...	...
45-60	...	...	...	...	...
60-75	...	...	...	...	...
75-90	...	...	...	...	...

3. Calcula a média dos tempos gastos diariamente a ver televisão pelos alunos inquiridos.

4. Constrói o histograma das frequências absolutas.

5. Indica a classe modal e localiza-a geometricamente.

6. Constrói o histograma das frequências relativas acumuladas.

7. Localiza geometricamente a mediana e os quartis e constrói o diagrama de extremos e quartis.

T21: Tarefa *RSI por família* (selecionada do manual adotado)

Supõe que, numa dada região, o Rendimento Social de Inserção por família obedece à seguinte distribuição:

Valor mensal (€)	N.º de famílias
[0 , 100[	6
[100 , 200[	19
[200 , 300[	34
[300 , 400[	6

Identifica a classe modal e diz o que significa no contexto da situação apresentada.

Determina geometricamente a moda da distribuição.

Determina a média dos valores mensais do Rendimento Social de Inserção atribuídos na região considerada.

Constrói um histograma de frequências relativas acumuladas e um diagrama de extremos e quartis desta distribuição.

T22: Tarefa *Reação positiva a um medicamento* (selecionada do manual adotado)

Os doentes que reagiram de forma positiva ao medicamento num período de tempo compreendido entre 40 e 50 minutos distribuem-se por idade e sexo da seguinte forma:

Homens		Mulheres
5 4 4 1 0	3	8 8 9 9 9
9 9 8 3 1	4	1 2 5 5 6 8
8 7 2 2	5	2 3 3 5 7
7 3 0 0	6	1 1 2 3 6 7 8
8 7 5	7	0 2 2 5 8 8

Determina a média, a moda e a mediana da distribuição das idades:

- do grupo das mulheres;
- do grupo dos homens;
- do grupo de doentes.

Para a distribuição das idades dos doentes de cada um dos sexos, constrói diagramas de extremos e quartis. Faz uma análise comparativa dos dois diagramas e apresenta algumas possíveis conclusões.

T23: Tarefa *Qual das duas gémeas a melhor aluna?* (criada pela professora e integrada numa ficha de trabalho)

A Joana e Mariana são irmãs gémeas. Estão no 10.º ano, na mesma turma, e gostam de competir pelos melhores resultados em Matemática. As suas notas nos testes de avaliação ao longo do ano letivo foram:

Joana	13	16	10	12	14	13	13
Mariana	13	13	8	18	13	7	19

Considera as notas da Joana a amostra 1 e as notas da Mariana a amostra 2.

a) Para cada uma das amostras, calcula a média, moda e mediana.

b) A média parece-te um bom indicador das notas das Gémeas?

c) O desempenho das duas foi semelhante?

NOTA: Vamos recorrer às chamadas **medidas de dispersão** para podermos tirar conclusões sobre os dados.

d) Calcula a amplitude e a amplitude interquartil de cada amostra. Compara-as. O que concluis?

e) Sem recorrer à calculadora, a não ser para efectuar cálculos, determina relativamente **às notas da Joana**, o valor da **variância** e do **desvio padrão**. **Para tal, segue as indicações:**

1º Preenche a seguinte tabela:

Notas Joana	Freq. Abs.	Desvio	Quadrado dos desvios	Desvio de cada valor x_i relativamente à média $\bar{x}$
x_i	n_i	$x_i - \bar{x}$	$\left(x_i - \bar{x}\right)^2$	$\left(x_i - \bar{x}\right)^2 n_i$
10	1	$10-13=-3$	$\left(-3\right)^2 = 9$	$9 \times 1 = 9$
12				
13				
14				
16				
				Soma total (*)

2º Divide o valor obtido em (*) pelo nº total de notas observadas.

Com este cálculo, acabaste de determinar o valor da **variância,** que é igual à média dos quadrados dos desvios e que pode se calculada diretamente pela fórmula:

$$\sigma^2 = \frac{\sum_{i=1}^{n} n_i \times \left(x_i - \bar{x}\right)^2}{N}$$

3º Calcula a raiz quadrada do valor obtido para a variância.

O valor obtido é o valor do **desvio-padrão,** que é uma das medidas estatísticas mais utilizadas para medir a variabilidade de um conjunto de dados quantitativos em relação à média, pode ser calculado através da fórmula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} n_i \times \left(x_i - \bar{x}\right)^2}{N}}$$

f) Procede da mesma forma relativamente às notas da Mariana, calcula valor da variância e do desvio padrão.

g) Compara os valores dos desvios-padrão da cada uma das amostras. O que concluis?

T24: Tarefa *Média e desvio-padrão em dados agrupados* (criada pela professora e integrada numa ficha de trabalho)

Observa os seguintes resultados dos testes de Inglês em duas turmas diferentes do 10ºano:

Turma A

Classificação	y_i	n_i
[0 , 5[		2
[5 , 10[		5
[10 , 15[		16
[15 , 20]		2

Turma B

Classificação	y_i	n_i
[0 , 5[		7
[5 , 10[		1
[10 , 15[		4
[15 , 20]		8

a) Completa as seguintes tabelas, calculando a marca da classe de cada turma - y_i .
b) Calcula valores aproximados da **média** e do **desvio-padrão** das notas dos testes em cada turma. Compara-os e interpreta-os.

T25: Tarefa *Medições na Eletrotecnia* (retirado do manual adotado).

Num trabalho prático de Electrotecnia, os alunos tiveram de medir um fio eléctrico.

Os resultados das medições, em centímetros, foram os seguintes:

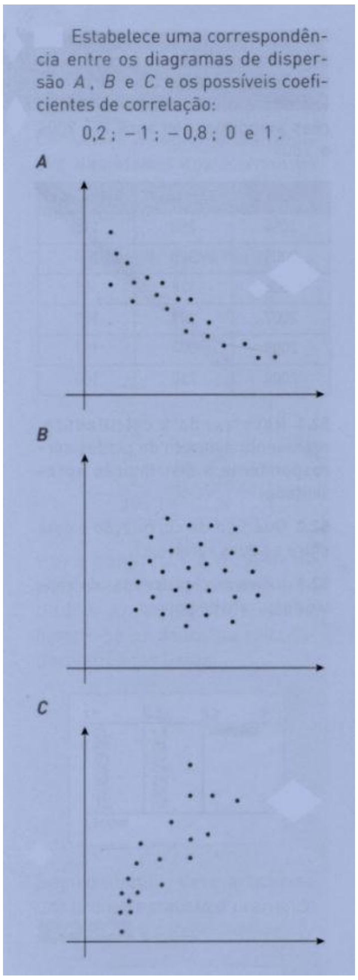
10,4	10,3	9,8	10,2
10	10,2	10,1	9,8
9,9	10	10,2	9,7

Qual foi a amplitude das medições?

Calcula a média, a variância e o desvio-padrão das medições efectuadas.

Compara os valores da média e da mediana. O que podes concluir?

T26: Tarefa *Associe entre diagramas de dispersão e coeficientes de correlação* (selecionada do manual adotado).


Estabelece uma correspondência entre os diagramas de dispersão A, B e C e os possíveis coeficientes de correlação:

$0{,}2$; -1 ; $-0{,}8$; 0 e 1.

T27: Tarefa *A evolução do crescimento de Pedro* (selecionada do manual adotado)

Na tabela seguinte estão representados alguns dados relativos à altura (cm) do Pedro. Recorre à calculadora e:

Idade (anos)	2	3	4	6	8	12	15
Altura (cm)	78	84	92	110	120	145	160

(1) Representa a nuvem de pontos associada aos dados.
(2) Determina a equação da reta de regressão e faz a respectiva representação gráfica.
(3) Determina os valores aproximados: da altura do pedro quando tinha 10 anos; da idade do Pedro quando a altura era 150 cm.

T28: Tarefa *Distribuição de médicos e farmácias por regiões* (selecionada do manual adotado)

O quadro seguinte refere-se à distribuição do número de médicos e de farmácias pelas diversas regiões do país, em 2007, com excepção dos dados referentes ao Alentejo, que não se encontram visíveis.

Referência dos dados (2007)	N.º de farmácias	N.º de médicos
Regiões	x_i	y_i
Norte	875	12 336
Centro	667	7 280
Lisboa	768	14 529
Alentejo		
Algarve	109	1 235
Região Autónoma dos Açores	47	475
Região Autónoma da Madeira	62	585

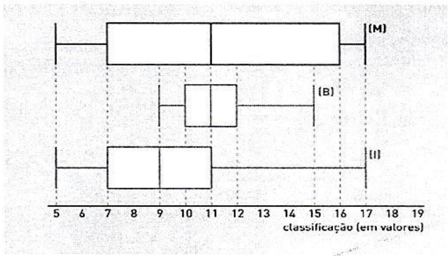
Fonte: INE

Recorre à calculadora gráfica para responderes às seguintes questões:

1. Constrói o diagrama de dispersão (x_i, y_i) e identifica o tipo de correlação existente.
2. Determina o coeficiente de correlação e traça a recta de regressão, escrevendo a respectiva equação da forma $y = ax + b$; $a, b \in \mathbb{R}$.
   Considera os valores de a e b arredondados às milésimas.
3. Faz uma estimativa do número de farmácias existentes na região do Alentejo, em 2007, sabendo que o número de médicos era de 1464.

T29: Tarefa *Compara o aproveitamento em três disciplinas dos mesmos alunos* (selecionada de um manual).

O André resolveu comparar o seu aproveitamento em três disciplinas: Inglês (I), Matemática (M) e Biologia (B). Ao longo do ano lectivo registou as classificações que obteve nas três disciplinas e apresentou os resultados em três diagramas de extremos e quartis:



Entre que valores variaram as classificações deste aluno, na disciplina de Biologia?

Em qual das disciplinas teve menos classificações negativas?

Entre que valores variaram 25% das classificações mais altas na disciplina de Matemática?

Entre que valores variaram 25% das classificações mais baixas na disciplina de Biologia?

Na tua opinião, em que disciplina o André obteve melhor aproveitamento? Justifica.

Situação *Como reduzir a taxa de absentismo na empresa*? (retirado de um manual e integrado num *powerpoint* exibido na aula)

Com o objetivo de reduzir a taxa de absentismo nos meses de inverno, o diretor geral de uma empresa pretende fazer uma sondagem para perceber o interesse dos empregados em receber gratuitamente a vacina contra a gripe. Para tal, recorre a uma amostra de 10% dos 200 funcionários da empresa. **Como selecionar a amostra?**

1. Uma hipótese é escolher os 20 funcionários sem obedecer a qualquer condição. Para isso, basta atribuir a cada funcionário um número de 1 a 200 e recorrer a um *software* para gerar aleatoriamente, vinte números (funcionários). Podemos recorrer, por exemplo, à calculadora gráfica. → **amostragem aleatória simples**

2. Outra hipótese é atribuir também a cada funcionário um número de 1 a 200 e a escolha ser feita seleccionando, por exemplo, o 9.º e a partir daí o 19.º, o 29.º, o 39.º, ... e assim sucessivamente, até encontrar os vinte números acabados em 9 que existem de 1 a 200. → **amostragem sistemática**

3. Ainda outra hipótese será dividir por estratos os funcionários da empresa. Por exemplo, por classes etárias. → **amostragem estratificada**

   Determina-se, proporcionalmente, a representação de cada uma das classes etárias e obtém-se o número desejado de pessoas a seleccionar:

Classe etária	Número de funcionários	Percentagem	Seleccionados
Menos de 30	50	25% (= 50/200)	25% × 20 = 5
Entre 30 e 50	140	70% (= 140/200)	70% × 20 = 14
Mais de 50	10	5% (= 10/200)	5% × 20 = 1
Total	200	100%	20

Serão então seleccionados aleatoriamente 5 funcionários com menos de 30 anos, 14 com idades entre 30 e 50 anos e 1 com mais de 50 anos.

Situação *Qual a amostra mais fiável?* (retirado de um manual e integrado num *powerpoint* exibido na aula)

A estação de televisão A pretende conhecer a preferência televisiva da população portuguesa. Para tal encomendou o estudo a três empresas distintas: Marks, Tests e Kmercs.

Foram apresentados os seguintes resultados:

Empresa Marks	
Estação A	80%
Estação B	1%
Estação C	7%
Estação D	12%

Ficha técnica: Esta sondagem foi realizada para a estação televisiva A, pela empresa Marks [...]. Foram feitas 1000 entrevistas telefónicas efectivas, das quais 500 a mulheres. Os números de telefone foram escolhidos aleatoriamente de entre os números das chamadas recebidas durante um programa televisivo da estação A.

Empresa Tests	
Estação A	31%
Estação B	15%
Estação C	26%
Estação D	28%

Ficha técnica: Esta sondagem foi realizada para a estação televisiva A, pela empresa Tests [...]. Foram feitas 1000 entrevistas telefónicas efectivas, das quais 650 a mulheres. Os números de telefone foram escolhidos aleatoriamente de uma lista telefónica abrangendo diferentes regiões do país.

Empresa Kmercs	
Estação A	30%
Estação B	13%
Estação C	19%
Estação D	38%

Ficha técnica: Esta sondagem foi realizada para a estação televisiva A, pela empresa Kmercs [...]. Foram feitas 100 entrevistas telefónicas efectivas, das quais 60 a mulheres. Os números de telefone foram escolhidos aleatoriamente de uma lista telefónica abrangendo diferentes regiões do país.

Como podes observar, os resultados apresentados são muito diferentes.

Situação *Zapping da Adriana* (retirado de um manual e integrado num *powerpoint* exibido na aula)

O pai da Adriana irrita-se com ela pois diz que ela não para de fazer *zapping* enquanto está a ver televisão. Para mostrar ao pai que não é verdade, a Adriana registou o número de vezes que fez *zapping* durante 31 dias do mês passado e calculou a média.

Classes	Frequência absoluta – f_i	Valor central da classe – x_i	$f_i\,x_i$
[0, 5[	9	2,5	22,5
[5, 10[	7	7,5	52,5
[10, 15[	4	12,5	50
[15, 20[	6	17,5	105
[20, 25[	5	22,5	112,5
Total	$\Sigma f_i = 31$		$\Sigma f_i\,x_i = 342{,}5$

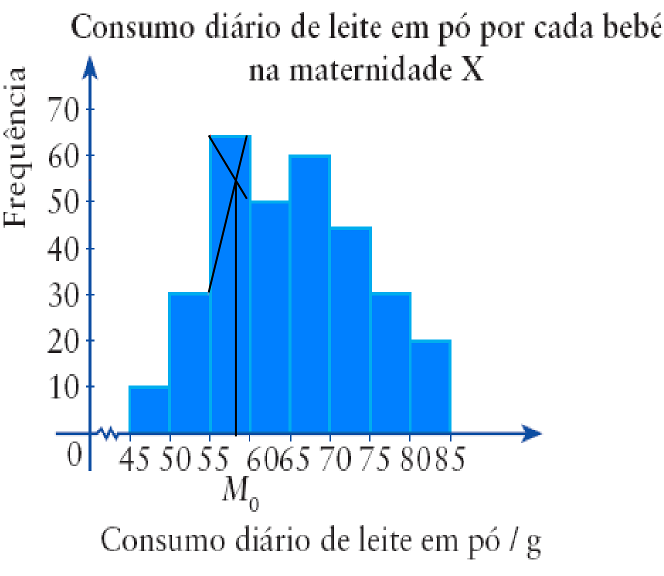
$$\overline{x} \approx \frac{342{,}5}{31}$$

$$\overline{x} \approx 11{,}05$$

Situação *Consumo diário de leite em pó numa maternidade* (retirado de um manual e integrado num *powerpoint* exibido na aula)

Numa maternidade, o consumo diário de leite em pó por cada bebé está registado na tabela seguinte. Qual a classe modal na tabela?

Leite (g)	N.° de bebés
[45 , 50[	11
[50 , 55[	31
[55 , 60[	65
[60 , 65[	48
[65 , 70[	60
[70 , 75[	46
[75 , 80[	30
[80 , 85[	21



ANEXO 5

A unidade de Estatística do Programa de Matemática A do ensino secundário (ME, 2001)

Tema III - Estatística

15 aulas de 90 minutos

Algumas das noções que se tratam nesta unidade já foram abordadas no 3º ciclo e, por isso, é possível em qualquer altura reinvestir nestes conhecimentos e completá-los progressivamente.
O estudante deverá ficar a saber organizar, representar e tratar dados recolhidos em bruto (ou tabelados) para daí tirar conclusões numa análise sempre crítica e sempre consciente dos limites do processo de matematização da situação. É importante que o estudo da Estatística contribua para melhorar a capacidade dos estudantes para avaliar afirmações de carácter estatístico, fornecendo-lhes ferramentas apropriadas para rejeitar quer certos anúncios publicitários quer notícias ou outras informações em que a interpretação de dados ou a realização da amostragem não tenha sido correcta.
Este tema fornece uma excelente oportunidade para actividades interdisciplinares, individualmente ou em grupo, devendo o professor ao definir o plano de trabalho com os estudantes incentivá-los a recorrer ao computador. No final, os estudantes devem interpretar e comunicar os resultados à turma fazendo uma análise crítica e estando conscientes que modos diferentes de apresentar as conclusões podem alterar a mensagem. No estudo deste tema o estudante deve recorrer à calculadora gráfica ou ao computador e às suas potencialidades para resolver muitos dos problemas.

Pré-Requisitos: Estatística do 3º ciclo do Ensino Básico.

Desenvolvimento	Indicações Metodológicas
Estatística – Generalidades ■ Objecto da Estatística e breve nota histórica sobre a evolução desta Ciência; utilidade na vida moderna. Clarificação de quais os fenómenos que podem ser objecto de estudo estatístico; exemplificação de tais fenómenos com situações da vida real, salientando o papel relevante da Estatística na sua descrição.	Deve-se chamar a atenção para o papel relevante desempenhado pela Estatística em todos os campos do conhecimento.

continua

Desenvolvimento	Indicações Metodológicas
■ Recenseamento e sondagem. As noções de população e amostra. Compreensão do conceito de amostragem e reconhecimento do seu papel nas conclusões estatísticas; distinção entre os estudos e conclusões sobre a amostra e a correspondente análise sobre a população. Noções intuitivas sobre as escolhas de amostras, sobre a necessidade de serem aleatórias, representativas e livres de vícios de concepção. ■ Estatística Descritiva e Estatística Indutiva.	Sendo a Estatística a Ciência que trata dos "dados", num procedimento estatístico estão envolvidas, de um modo geral, duas fases: uma fase de organização dos dados recolhidos, em que se procura reduzir, de forma adequada, a informação neles contida - **Estatística Descritiva**, e uma segunda fase, em que se procura tirar conclusões e tomar decisões para um conjunto mais vasto, de onde se recolheram os dados - **Inferência Estatística**. Existe, no entanto, uma fase pioneira, que diz respeito à aquisição dos próprios "dados". Deve-se realçar a importância de, ao iniciar qualquer estudo estatístico, proceder cuidadosamente ao planeamento da experiência que conduz à recolha dos "dados" que serão objecto de tratamento estatístico.
Organização e interpretação de caracteres estatísticos (qualitativos e quantitativos) ■ Análise gráfica de atributos qualitativos (gráficos circulares, diagramas de barras, pictogramas); determinação da moda; ■ Análise de atributos quantitativos: variável discreta e variável contínua. Dados agrupados em classes. ■ Variável discreta; função cumulativa. ■ Variável contínua: tabelas de frequências (absolutas, relativas e relativas acumuladas); gráficos (histograma, polígono de frequências); função cumulativa. ■ Medidas de localização de uma amostra: moda ou classe modal; média; mediana; quartis.	Deve-se chamar a atenção para o facto de que a organização dos dados, consiste em resumir a informação neles contida através de tabelas, gráficos e algumas medidas, a que damos o nome de "estatísticas". Nesta fase, em que se substitui todo o conjunto dos dados, por um sumário desses dados, devem-se tomar as devidas precauções, pois nem todos os instrumentos de redução de dados se aplicam a todos os tipos de dados. Assim, de entre esses processos deve-se ter presente quais os mais adequados e em que situações é ou não convenientes aplicá-los. A título de exemplo referimos o facto de não ter qualquer sentido calcular a média para dados de tipo qualitativo, mesmo que as diferentes categorias assumidas pela variável em estudo estejam representadas por números.

continua

Desenvolvimento	Indicações Metodológicas
■ Medidas de dispersão de uma amostra: amplitude; variância; desvio padrão; amplitude interquartis. ■ Discussão das limitações destas estatísticas. ■ Diagramas de "extremos e quartis"	
Referência a distribuições bidimensionais (abordagem gráfica e intuitiva) ▪ Diagrama de dispersão; dependência estatística; ideia intuitiva de correlação; exemplos gráficos de correlação positiva, negativa ou nula. ▪ Coeficiente de correlação e sua variação em $[-1, 1]$. ▪ Definição de centro de gravidade de um conjunto finito de pontos; sua interpretação física. ▪ Ideia intuitiva de recta de regressão; sua interpretação e limitações.	Generalizando o estudo de uma única variável, faz-se uma introdução ao estudo dos dados bivariados, insistindo na representação gráfica sob a forma do diagrama de dispersão ou diagrama de pontos. Quando, a partir desta representação, se verificar uma tendência para a existência de uma associação linear entre as duas variáveis em estudo, identifica-se uma medida que quantifica o grau de associação - o coeficiente de correlação, assim como se apresenta um modelo matemático que permitirá, conhecido o valor de uma das variáveis, obter uma estimativa para o valor da outra variável.

ANEXO 6

Pedidos de autorização

Exmo. Sr. Diretor da Escola Secundária ____________________________

Eu, Sandra Maria Oliveira Quintas, professora do grupo 500, venho por este meio solicitar autorização para concretizar nesta escola um Projeto de Investigação em Educação sobre o conhecimento didático do professor em Estatística. Este projeto integra-se no âmbito do Doutoramento em Educação, na área de especialização em Didática da Matemática, do Instituto de Educação da Universidade de Lisboa, que estou a desenvolver e tem como objetivo principal analisar e compreender o conhecimento didático do professor em Estatística quando ensina os diferentes tópicos estatísticos, na disciplina de Matemática A, do ensino secundário.

A concretização deste projeto implicará a recolha de dados no 3º período do presente ano letivo, acompanhando professora e alunos do 10º ano, ocorrendo em momentos em que o tema da Estatística é lecionado. Solicito que a recolha de dados seja feita na turma de 10.º ano (Matemática A) lecionada pela professora de Matemática Dr.ª _____________que se mostrou disponível para ajudar a viabilizar este projeto. Convém referir que, deste trabalho, não resultará qualquer prejuízo nem para a professora nem para os alunos, não havendo lugar a qualquer alteração relativamente aos conteúdos programados nem às indicações metodológicas referidas no programa de Matemática A em vigor, sendo a professora da turma, em conjunto com o seu grupo disciplinar, responsáveis pela planificações e lecionação do tema.

As principais formas de recolha de dados para a concretização deste projecto serão: i) entrevista; ii) observação participante; iii) materiais produzidos ou utilizados pela professora na lecionação do tema da estatística; iv) produções escritas dos alunos na resolução de algumas tarefas. A recolha de dados envolverá o registo áudio e/ou vídeo de alguns destes momentos, uma vez colhida a anuência dos alunos. As entrevistas decorrerão, caso me seja autorizada a realização deste projecto nesta escola, nas instalações da própria escola e em horários a combinar com a professora.

Em todo o processo serão salvaguardados os direitos de privacidade e anonimato que assistem aos participantes e à própria escola, enquanto instituição. Os encarregados de educação serão informados sobre este estudo, sendo essencial o seu consentimento para possibilitar a participação dos alunos neste projeto de investigação.

_____________________, 22 de Abril de 2011

Pede deferimento,

____

(Sandra Quintas)

Exmo(a). Sr(a). Encarregado(a) de Educação

Eu, Sandra Maria Oliveira Quintas, professora do grupo 500, atual formadora, com certificado emitido pelo Conselho Científico-Pedagógico da Formação Contínua, no Centro de formação____, venho por este meio solicitar autorização para a participação do seu educando num projeto de investigação em Educação sobre o conhecimento didático do professor em Estatística. Este projeto integra-se no âmbito do Doutoramento em Educação, na área de especialização em Didática da Matemática, do Instituto de Educação da Universidade de Lisboa, que estou a desenvolver e tem como objetivo principal analisar e compreender o conhecimento didático do professor em Estatística quando ensina os diferentes tópicos estatísticos, na disciplina de Matemática A, do ensino secundário.

A professora da disciplina de Matemática, Dr.ª __________________________, mostrou a sua total disponibilidade para viabilizar este projeto. A concretização deste projeto implicará a recolha de dados, na sala de aula, durante o 3º período, usando meios áudio e vídeo, acompanhando a professora e alunos do 10º ano e ocorrerá em momentos em que o tema da Estatística é lecionado. De modo a obter informação mais detalhada está ainda prevista a recolha de algumas resoluções escritas dos alunos. As imagens e o som daí resultantes não serão divulgados nem serão utilizados para quaisquer outros fins, senão para este trabalho académico, sendo sempre preservado o anonimato, quer dos alunos e da professora, quer da escola. Procederei à destruição destas imagens após conclusão deste trabalho.

Convém referir que, deste trabalho, não resultará qualquer prejuízo para os alunos, não havendo lugar a qualquer alteração relativamente aos conteúdos programados nem às indicações metodológicas referidas no programa de Matemática A em vigor, nem este terá quaisquer repercussões na avaliação dos alunos. Considero que, a participação dos alunos no estudo poderá revelar-se uma mais-valia para a sua aprendizagem matemática.

Para quaisquer esclarecimentos adicionais, poderá contactar-me através do e-mail: sandramquintas@hotmail.com

Com os meus melhores cumprimentos.

___________________, 22 de Abril de 2011

A Investigadora

____

(Sandra Quintas)

Autorização

Eu,____, Encarregado(a) de Educação do(a) aluno(a)______________________________________, Nº_____, da turma_____, do 10.º ano, da Escola Secundária____, declaro que tomei conhecimento dos objetivos do projeto de investigação sobre o conhecimento didático do professor do ensino secundário em Estatística e da necessidade de alguns momentos de recolha de dados serem áudio e vídeo gravados.

Neste sentido autorizo/não autorizo **(riscar o que não interessa)** a participação do meu educando no estudo, desde que seja salvaguardado o seu anonimato, seja protegida a sua privacidade e sejam destruídos todos os vídeos e gravações áudio após a conclusão deste trabalho.

_____/_____/_____ (data)

O(a) Encarregado(a) de Educação

____